GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:50:34 ; Search time 621.5 Seconds
(without alignments)
4354.885 Million cell updates/sec

Title: US-09-091-605-2
Perfect score: 93
Sequence: 1 CATGCTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	93	100.0	396	6	AX147675	AX147675 Sequence
2	93	100.0	528	6	AR030615	AR030615 Sequence
3	93	100.0	528	6	AR168153	AR168153 Sequence
4	93	100.0	528	6	E05860	E05860 DNA encodin
5	93	100.0	559	4	AF529185	AF529185 Ovis arie
6	93	100.0	955	6	AR108107	AR108107 Sequence
7	93	100.0	955	6	AR108109	AR108109 Sequence
8	93	100.0	1062	9	HUMGLUC	J04040 Human gluca
9	93	100.0	1108	4	BOVGG	K00107 Bovine panc
10	93	100.0	1154	9	BC005278	BC005278 Homo sapi
11	93	100.0	2356	6	AR108119	AR108119 Sequence
12	93	100.0	6455	9	HSGLUC	V01515 Human gene
13	93	100.0	10050	9	HSGLUCG2	X03991 Human gluca
c 14	93	100.0	163681	9	AC007750	AC007750 Homo sapi
15	91.4	98.3	528	6	AR030614	AR030614 Sequence
16	91.4	98.3	528	6	AR168152	AR168152 Sequence
17	91.4	98.3	1123	4	AF308439	AF308439 Canis fam
18	88.8	95.5	1053	10	GPIGG	D00014 Cavia porce
19	86.6	93.1	1104	10	OCOGLU	M57688 Octodon deg
20	83.4	89.7	277	10	RATGLU4	K02811 Rat glucago
21	83.4	89.7	895	6	AR108106	AR108106 Sequence
22	83.4	89.7	1034	6	A31421	A31421 H.sapiens m
23	83.4	89.7	1102	10	BC012975	BC012975 Mus muscu
c 24	83.4	89.7	182906	2	AC111919	AC111919 Rattus no
c 25	83.4	89.7	210317	2	AC115074	AC115074 Mus muscu
26	83.4	89.7	219014	2	AC024141	AC024141 Mus muscu
27	81.8	88.0	668	10	MMPPROGLG	Z46845 M.musculus
28	81.8	88.0	1116	10	AF276754	AF276754 Mus muscu
29	80.8	86.9	144	6	AR030619	AR030619 Sequence
30	80.8	86.9	144	6	AR168157	AR168157 Sequence
31	79	84.9	106	11	G42680	G42680 GCG Bovine
32	78.6	84.5	1118	10	HAMGG	J00059 Syrian hams
33	72.2	77.6	1186	5	S78477	S78477 proglucagon
34	72.2	77.6	1576	5	GGPPG	Y07539 Chicken mRN
35	72.2	77.6	1576	6	E03593	E03593 DNA encodin
36	65.8	70.8	1050	5	HSU77612	U77612 Heloderma s
37	65.8	70.8	1196	5	HSU77611	U77611 Heloderma s
38	63.2	68.0	315	6	I07551	I07551 Sequence 40
39	50	53.8	1408	5	AF004432	AF004432 Xenopus l
40	49.8	53.5	78	6	AR077755	AR077755 Sequence
41	47.6	51.2	1290	5	AF004433	AF004433 Xenopus l
42	46	49.5	1300	5	AF324209	AF324209 Hoplobatr
43	45.6	49.0	493	5	LAGL01	V00632 Lophius ame
44	41.4	44.5	383	5	OMU19913	U19913 Oncorhynchu
45	41.4	44.5	383	5	OTU19920	U19920 Oncorhynchu

ALIGNMENTS

RESULT 1
AX147675
LOCUS AX147675 396 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 2 from Patent WO0136643.
ACCESSION AX147675
VERSION AX147675.1 GI:14346730
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 396)
AUTHORS Treco,D.A., Concino,M.F. and Duguay,S.J.
TITLE Nucleic acid construct for optimized production of products
JOURNAL Patent: WO 0136643-A 2 25-MAY-2001;
TRANSKARYOTIC THERAPIES, INC. (US)

```
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthesized nucleotide"
BASE COUNT       81 a    111 c    121 g     83 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 6;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     286 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 345

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     346 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 378


RESULT 2
AR030615
LOCUS       AR030615                528 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 8 from patent US 5861284.
ACCESSION   AR030615
VERSION     AR030615.1  GI:5943829
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,O., Kuriyama,M., Koyama,N. and Fukuda,T.
  TITLE     Method for producing a biologically active recombinant
            cysteine-free parathyroid hormone (1-34)
  JOURNAL   Patent: US 5861284-A 8 19-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="unknown"
BASE COUNT      142 a    121 c    143 g    122 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 6;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 3
AR168153
LOCUS       AR168153                528 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 8 from patent US 6287806.
ACCESSION   AR168153
VERSION     AR168153.1  GI:17903977
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,O., Kuriyama,M., Koyama,N. and Fukuda,T.
  TITLE     Method for producing a biologically active recombinant
            cysteine-free peptide
  JOURNAL   Patent: US 6287806-A 8 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="unknown"
BASE COUNT      142 a    121 c    143 g    122 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 6;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 4
E05860
LOCUS       E05860                  528 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  DNA encoding fusion protein comprising glucogon-like peptide 1 and
            fibroblast growth factor.
ACCESSION   E05860
VERSION     E05860.1  GI:2174047
KEYWORDS    JP 1993304976-A/3.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,T., Koyama,N., Kuriyama,M. and Fukuda,T.
  TITLE     PRODUCTION OF PEPTIDE
  JOURNAL   Patent: JP 1993304976-A 3 19-NOV-1993;
            TAKEDA CHEM IND LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            OS   Homo sapiens (human)
            PN   JP 1993304976-A/3
            PD   19-NOV-1993
            PF   18-FEB-1992 JP 1992030635
            PR   19-FEB-1991 JP 91P    24841, 18-OCT-1991 JP 91P   271438, PR
            24-OCT-1991 JP 91P   277724
            PI   NISHIMURA TADASHI, KOYAMA NOBUYUKI, KURIYAMA MASATO, PI
            FUKUDA TSUNEHIKO
            PC   C12P21/00,C07K15/00,C12N15/00//C12N15/18,C12N15/62,C12P21/02,
            PC   (C12P21/00,
            PC   C12R1:19),(C12P21/02,C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   CDS             1. .528
            FT                   /product='fusion protein comprising internal
            FT                   fragment of
            FT                   glucogon-like peptide 1 and partial fragment
            FT                   of fibroblast
            FT                   growth factor'
            FT   mutation        replace(96,'T')
            FT                   /product='fusion comprising internal fragment
            FT                   of
            FT                   glucogon-like peptide 1 and partial fragment
            FT                   of fibroblast
            FT                   growth factor'
            FT   mat_peptide     1. .96
            FT                   /product='internal fragment of glucogon-like
            FT                   peptide 1'
            FT   mutation        replace(96,'T')
            FT                   /product='internal fragment of glucogon-like
            FT                   peptide 1'
            FT   mat_peptide     97. .528
            FT                   /product='partial fragment of fibroblast  FT
                              growth factor'.
FEATURES             Location/Qualifiers
```

```
     source          1. .528
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      142 a    122 c    143 g    121 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 6;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 5
AF529185
LOCUS       AF529185                559 bp    mRNA    linear   MAM 13-AUG-2002
DEFINITION  Ovis aries preproglucagon mRNA, partial cds.
ACCESSION   AF529185
VERSION     AF529185.1  GI:22212831
KEYWORDS    .
SOURCE      sheep.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Limesand,S.W. and Hay,W.W. Jr.
  TITLE     Characterization of the endocrine pancreas in an ovine placental
            insufficiency IUGR fetus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 559)
  AUTHORS   Limesand,S.W. and Hay,W.W. Jr.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2002) Pediatrics, University of Colorado Health
            Sciences Center, 4200 E 9th Ave, Denver, CO 80262, USA
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
                     /tissue_type="pancreas"
                     /dev_stage="fetal"
     5'UTR           1. .30
     CDS             31. .>559
                     /codon_start=1
                     /product="preproglucagon"
                     /protein_id="AAM94409.1"
                     /db_xref="GI:22212832"
                     /translation="MKSLYFVAGLLVMLAQGSWQHSLQNTEEKSSSFPAPQTDPLGDP
                     DQISEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVNIVEELRRRHADGSFSDEMNTVLD
                     SLATRDFINWLLQTKI"
BASE COUNT      160 a    137 c    138 g    124 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 4;  Length 559;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 381

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    382 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 414


RESULT 6
AR108107
LOCUS       AR108107                955 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 57 from patent US 6110707.
ACCESSION   AR108107
VERSION     AR108107.1  GI:12823594
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 955)
  AUTHORS   Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
            Quaade,C., Kruse,F. and McGarry,D.
  TITLE     Recombinant expression of proteins from secretory cell lines
  JOURNAL   Patent: US 6110707-A 57 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .955
                     /organism="unknown"
BASE COUNT      301 a    181 c    203 g    270 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 6;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 7
AR108109
LOCUS       AR108109                955 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 60 from patent US 6110707.
ACCESSION   AR108109
VERSION     AR108109.1  GI:12823596
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 955)
  AUTHORS   Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
            Quaade,C., Kruse,F. and McGarry,D.
  TITLE     Recombinant expression of proteins from secretory cell lines
  JOURNAL   Patent: US 6110707-A 60 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .955
                     /organism="unknown"
BASE COUNT      302 a    180 c    202 g    271 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 6;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 8
HUMGLUC
LOCUS       HUMGLUC                 1062 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human glucagon mRNA, complete cds.
ACCESSION   J04040
```

```
VERSION     J04040.1  GI:183269
KEYWORDS    glicentin; glucagon.
SOURCE      Human (neonate) brainstem, cDNA to mRNA, clones BS[3,A,B].
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1062)
  AUTHORS   Drucker,D.J. and Asa,S.
  TITLE     Glucagon gene expression in vertebrate brain
  JOURNAL   J. Biol. Chem. 263 (27), 13475-13478 (1988)
  MEDLINE   88330860
   PUBMED   2901414
FEATURES             Location/Qualifiers
     source          1. .1062
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="2q36-q37"
     gene            1. .1062
                     /gene="GCG"
     mRNA            <1. .1062
                     /gene="GCG"
                     /product="GCG mRNA"
     CDS             38. .580
                     /gene="GCG"
                     /note="preproglucagon"
                     /codon_start=1
                     /protein_id="AAA52567.1"
                     /db_xref="GI:183270"
                     /db_xref="GDB:G00-119-265"
                     /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                     DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                     NLAARDFINWLIQTKITDRK"
     sig_peptide     38. .97
                     /gene="GCG"
                     /note="glucagon signal peptide"
     mat_peptide     98. .304
                     /gene="GCG"
                     /product="glicentin"
     mat_peptide     194. .280
                     /gene="GCG"
                     /product="glucagon"
     mat_peptide     287. .304
                     /gene="GCG"
                     /product="intervening peptide I"
     mat_peptide     311. .421
                     /gene="GCG"
                     /product="glucagon-like peptide I"
     mat_peptide     428. .466
                     /gene="GCG"
                     /product="intervening peptide II"
     mat_peptide     473. .571
                     /gene="GCG"
                     /product="glucagon-like peptide II"
BASE COUNT      340 a    200 c    215 g    307 t
ORIGIN      81 bp upstream of RsaI site; chromosome 2q36-q37.

  Query Match            100.0%;  Score 93;  DB 9;  Length 1062;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    329 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 388

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    389 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 421


RESULT 9
BOVGG
LOCUS       BOVGG                   1108 bp    mRNA    linear   MAM 27-APR-1993
DEFINITION  Bovine pancreas preproglucagon mRNA.
ACCESSION   K00107
VERSION     K00107.1  GI:163081
KEYWORDS    glicentin; glucagon; hormone.
SOURCE      Bovine cdna to pancreatic mrna.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 1108)
  AUTHORS   Lopez,L.C., Frazier,M.L., Su,C.J., Kumar,A. and Saunders,G.F.
  TITLE     Mammalian pancreatic preproglucagon contains three glucagon-related
            peptides
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 80 (18), 5485-5489 (1983)
  MEDLINE   83299996
   PUBMED   6577439
FEATURES             Location/Qualifiers
     source          1. .1108
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
     CDS             91. .633
                     /note="preproglucagon"
                     /codon_start=1
                     /protein_id="AAA30538.1"
                     /db_xref="GI:163082"
                     /translation="MKSLYFVAGLFVMLVQGSWQRSLQNTEEKSSSFPAPQTDPLGDP
                     DQINEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVNIVEELRRRHADGSFSDEMNTVLD
                     SLATRDFINWLLQTKITDRK"
BASE COUNT      356 a    232 c    213 g    307 t
ORIGIN

  Query Match            100.0%;  Score 93;  DB 4;  Length 1108;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    382 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 441

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    442 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 474


RESULT 10
BC005278
LOCUS       BC005278                1154 bp    mRNA    linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, glucagon, clone MGC:12325 IMAGE:3950435, mRNA,
            complete cds.
ACCESSION   BC005278
VERSION     BC005278.1  GI:13528971
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1154)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
```

```
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 16 Row: k Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4503944.
FEATURES             Location/Qualifiers
     source          1. .1154
                     /organism="Homo sapiens"
                     /db_xref="LocusID:2641"
                     /db_xref="taxon:9606"
                     /clone="MGC:12325 IMAGE:3950435"
                     /tissue_type="Pancreas"
                     /clone_lib="NIH_MGC_78"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     CDS             100. .642
                     /codon_start=1
                     /product="glucagon"
                     /protein_id="AAH05278.1"
                     /db_xref="GI:13528972"
                     /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                     DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                     NLAARDFINWLIQTKITDRK"
BASE COUNT      391 a    218 c    226 g    319 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 9;  Length 1154;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    391 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 450

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    451 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 483


RESULT 11
AR108119
LOCUS       AR108119                2356 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 72 from patent US 6110707.
ACCESSION   AR108119
VERSION     AR108119.1  GI:12823606
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2356)
  AUTHORS   Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
            Quaade,C., Kruse,F. and McGarry,D.
  TITLE     Recombinant expression of proteins from secretory cell lines
  JOURNAL   Patent: US 6110707-A 72 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .2356
                     /organism="unknown"
BASE COUNT      614 a    600 c    581 g    561 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 6;  Length 2356;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1704 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 1763

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db   1764 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 1796


RESULT 12
HSGLUC
LOCUS       HSGLUC                  6455 bp    DNA     linear   PRI 09-FEB-1999
DEFINITION  Human gene encoding preproglucagon. Glucagon is a 29-amino acid
            pancreatic hormone which counteracts the blood glucose-lowering
            action of insulin by stimulating hepatic glycogenolysis and
            gluconeogenesis. Also included in the proglucagon sequence are two
            regions (GLP-1 and GLP-2) which are homologous to glucagon itself
            but not identical.
ACCESSION   V01515
VERSION     V01515.1  GI:31777
KEYWORDS    glucagon; signal peptide.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6455)
  AUTHORS   Bell,G.I., Sanchez-Pescador,R., Laybourn,P.J. and Najarian,R.C.
  TITLE     Exon duplication and divergence in the human preproglucagon gene
  JOURNAL   Nature 304 (5924), 368-371 (1983)
  MEDLINE   83271477
   PUBMED   6877358
COMMENT     Data kindly reviewed (20-SEP-1983) by G.I. Bell.
FEATURES             Location/Qualifiers
     source          1. .6455
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             join(168. .259,1832. .1993,3661. .3798,5168. .5315)
                     /codon_start=1
                     /product="precursor"
                     /protein_id="CAA24759.1"
                     /db_xref="GI:31778"
                     /db_xref="SWISS-PROT:P01275"
                     /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                     DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                     NLAARDFINWLIQTKITDR"
     sig_peptide     168. .227
     mat_peptide     1896. .1982
                     /product="glucagon"
     mat_peptide     3680. .3790
                     /product="GLP-1"
     mat_peptide     5211. .5312
                     /product="GLP-2"
     intron          260. .1831
                     /number=1
     intron          1994. .3660
                     /number=2
     intron          3799. .5167
                     /number=3
BASE COUNT     2167 a   1107 c   1121 g   2060 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 9;  Length 6455;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3698 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 3757

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db   3758 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 3790


RESULT 13
HSGLUCG2
```

```
LOCUS       HSGLUCG2              10050 bp    DNA     linear   PRI 20-APR-1995
DEFINITION  Human glucagon gene.
ACCESSION   X03991
VERSION     X03991.1  GI:31786
KEYWORDS    glucagon; glucagon-like peptide; hormone; preproglucagon;
            preprohormone; signal peptide.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10050)
  AUTHORS   White,J.W. and Saunders,G.F.
  TITLE     Structure of the human glucagon gene
  JOURNAL   Nucleic Acids Res. 14 (12), 4719-4730 (1986)
  MEDLINE   86259053
   PUBMED   3725587
COMMENT     See also <HSGLUC; V01515> which is derived from the same library.
            Sequence discrepancies may indicate that the two clones represent
            different alleles or may be due to cloning artefacts. pos. 3516 in
            <X03991> corresponds to pos. 9 in <V01515> (the first eight
            nucleotides of which are probably a cloning artefact).
FEATURES             Location/Qualifiers
     source          1. .10050
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="lambdahGCG1"
                     /clone_lib="fetal liver genomic DNA"
     promoter        complement(535. .538)
                     /note="CAAT-box (complementary strand)"
     promoter        577. .581
                     /note="TATA-box"
     exon            603. .698
                     /note="exon 1; untranslated leader (5' UT-region)"
     intron          699. .3665
                     /note="intron I"
     exon            3666. .3766
                     /note="exon 2"
     CDS             join(3675. .3766,5339. .5500,7177. .7314,8683. .8826,
                     9481. .9487)
                     /note="prepro-glucagon"
                     /codon_start=1
                     /protein_id="CAA27627.1"
                     /db_xref="GI:762941"
                     /db_xref="SWISS-PROT:P01275"
                     /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                     DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTNRNRNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                     NLAARDFINWLIQTKITDRK"
     sig_peptide     3675. .3734
     intron          3767. .5338
                     /note="intron II"
     exon            5339. .5500
                     /note="exon 3"
     misc_feature    5339. .5396
                     /note="glicentin related pancreatic peptide (AA 12-30)
                     (5339 is 3rd base in codon)"
     misc_feature    5397. .5402
                     /note="spacer (AA 31-32; 2AA)"
     misc_feature    5403. .5489
                     /note="glucagon (AA 33-61; 29AA)"
     misc_feature    5490. .5500
                     /note="spacer (AA 61-65; 4AA) (5500 is 2nd base in codon)"
     intron          5501. .7176
                     /note="intron III"
     exon            7177. .7314
                     /note="exon 4"
     misc_feature    7177. .7195
                     /note="spacer (AA 66-71; 6AA) (7177 is 3rd base in codon)"
     misc_feature    7196. .7306
                     /note="glucagon-like peptide 1 (AA 72-108; 37AA)"
     misc_feature    7307. .7314
                     /note="spacer (AA 109-111; 3AA) (7314 is 2nd base in
                     codon)"
     intron          7315. .8682
                     /note="intron IV"
     exon            8683. .8826
                     /note="exon 5"
     misc_feature    8683. .8826
                     /note="pro-glucagon (AA 112-159) (8683 is 3rd base in
                     codon) (8826 is 2nd base in codon)"
     misc_feature    8683. .8725
                     /note="spacer (AA 112-125; 14AA) (8683 is 3rd base in
                     codon)"
     misc_feature    8726. .8826
                     /note="glucagon-like peptide 2 (AA 126-159) (8826 is 2nd
                     base in codon)"
     intron          8827. .9480
                     /note="intron V"
     exon            9481. .>9487
                     /note="exon 6"
     misc_feature    9481. .9484
                     /note="glucagon-like peptide 2 (AA 159-160; 35AA total)
                     (9481 is 3rd base in codon)"
     misc_feature    9481. .9484
                     /note="pro-glucagon (AA 159-160) (9481 is 3rd base in
                     codon)"
     misc_feature    9950. .9956
                     /note="pot. polyA signal"
     misc_feature    9965. .9974
                     /note="region of pot. polyA sites"
BASE COUNT     3397 a   1698 c   1746 g   3209 t
ORIGIN

  Query Match           100.0%;  Score 93;  DB 9;  Length 10050;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7214 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 7273

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    7274 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 7306


RESULT 14
AC007750/c
LOCUS       AC007750              163681 bp    DNA     linear   PRI 02-OCT-2000
DEFINITION  Homo sapiens BAC clone RP11-576I16 from 2, complete sequence.
ACCESSION   AC007750
VERSION     AC007750.3  GI:6094634
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163681)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 163681)
  AUTHORS   Cotton,M., Maupin,R., Hawkins,M. and Harkins,R.
  TITLE     The sequence of Homo sapiens BAC clone RP11-576I16
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 163681)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 163681)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (22-OCT-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 163681)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Oct 22, 1999 this sequence version replaced gi:5103889.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0576I16
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-178A14.  Actual start of
            this clone is at base position 1 of RP11-576I16; actual end is at
            base position 163681 of RP11-576I16.
FEATURES             Location/Qualifiers
     source          1. .163681
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-576I16"
                     /clone_lib="RPCI-11"
     repeat_region   216. .322
                     /rpt_family="MER81"
     repeat_region   1271. .1298
                     /rpt_family="(CAGCT)n"
     repeat_region   1536. .1694
                     /rpt_family="MIR"
     repeat_region   2098. .3518
                     /rpt_family="L2"
     repeat_region   3614. .3902
                     /rpt_family="Alu"
     repeat_region   4090. .4568
                     /rpt_family="MER1_type"
     repeat_region   5128. .5148
                     /rpt_family="AT_rich"
     repeat_region   5263. .5299
                     /rpt_family="(CA)n"
     repeat_region   5417. .5648
                     /rpt_family="L1"
     repeat_region   7093. .7556
                     /rpt_family="L2"
     repeat_region   8702. .9187
                     /rpt_family="MaLR"
     repeat_region   9386. .9586
                     /rpt_family="MIR"
     repeat_region   10580. .10739
                     /rpt_family="L2"
     repeat_region   10792. .11064
                     /rpt_family="Alu"
     repeat_region   11136. .12912
                     /rpt_family="L1"
     misc_feature    12894. .12899
                     /note="match to EST AI077564 (NID:g3411972) oz33g05.x1"
     repeat_region   12982. .13011
                     /rpt_family="(TTG)n"
     repeat_region   13109. .13133
                     /rpt_family="(TTA)n"
     repeat_region   13134. .13419
                     /rpt_family="Alu"
     repeat_region   13833. .14130
                     /rpt_family="Alu"
     repeat_region   14437. .14573
                     /rpt_family="MER1_type"
     repeat_region   14742. .14887
                     /rpt_family="L1"
     repeat_region   15098. .15234
                     /rpt_family="L1"
     misc_feature    15444. .15460
                     /note="match to EST AI922427 (NID:g5658391) wo06b10.x1"
     repeat_region   15449. .15475
                     /rpt_family="(T)n"
     repeat_region   15476. .15695
                     /rpt_family="L1"
     repeat_region   15713. .15993
                     /rpt_family="Alu"
     repeat_region   15994. .16013
                     /rpt_family="(TAAA)n"
     repeat_region   16039. .16131
                     /rpt_family="L1"
     repeat_region   16140. .16281
                     /rpt_family="L1"
     repeat_region   16284. .16327
                     /rpt_family="(TG)n"
     repeat_region   16328. .16394
                     /rpt_family="(CATA)n"
     repeat_region   16396. .16576
                     /rpt_family="L1"
     repeat_region   16625. .16997
                     /rpt_family="L1"
     repeat_region   17020. .17274
                     /rpt_family="L1"
     repeat_region   17270. .17603
                     /rpt_family="L1"
     repeat_region   17644. .18442
                     /rpt_family="L1"
     repeat_region   18444. .20187
                     /rpt_family="L1"
     repeat_region   20260. .20562
                     /rpt_family="MER2_type"
     repeat_region   20563. .20850
                     /rpt_family="Alu"
     repeat_region   20851. .20910
                     /rpt_family="MER2_type"
     repeat_region   21194. .21218
                     /rpt_family="(T)n"
     misc_feature    21204. .21219
                     /note="match to EST AI582416 (NID:g4568313) tr97d11.x1"
     misc_feature    21204. .21219
```

```
                /note="match to EST AI583407 (NID:g4569304) ts09d01.x1"
     misc_feature    21204. .21219
                /note="match to EST AI589439 (NID:g4598487) tr76a08.x1"
     misc_feature    21204. .21219
                /note="match to EST AI865901 (NID:g5530008) wk88b06.x1"
     repeat_region   21618. .21730
                /rpt_family="MER1_type"
     repeat_region   21800. .21840
                /rpt_family="AT_rich"
     repeat_region   21991. .22290
                /rpt_family="Alu"
     misc_feature    21991. .22005
                /note="match to EST AI682812 (NID:g4892994) wc66f03.x1"
     repeat_region   22464. .22658
                /rpt_family="MaLR"
     repeat_region   22756. .23030
                /rpt_family="MaLR"
     repeat_region   23343. .23397
                /rpt_family="(CA)n"
     repeat_region   23919. .23983
                /rpt_family="MIR"
     repeat_region   24424. .24581
                /rpt_family="MER1_type"
     misc_feature    24728. .24983
                /note="match to EST T10929 (NID:g391083)"
     misc_feature    24730. .25234

  Query Match            100.0%;  Score 93;  DB 9;  Length 163681;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27495 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 27436

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db  27435 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 27403


RESULT 15
AR030614
LOCUS       AR030614                 528 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 7 from patent US 5861284.
ACCESSION   AR030614
VERSION     AR030614.1  GI:5943828
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,O., Kuriyama,M., Koyama,N. and Fukuda,T.
  TITLE     Method for producing a biologically active recombinant
            cysteine-free parathyroid hormone (1-34)
  JOURNAL   Patent: US 5861284-A 7 19-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="unknown"
BASE COUNT      142 a    121 c    144 g    121 t
ORIGIN

  Query Match            98.3%;  Score 91.4;  DB 6;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 4.3e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACGAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
```

```
Search completed: February 14, 2003, 08:24:43
Job time : 665.5 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:57:09 ; Search time 1184 Seconds
(without alignments)
1272.112 Million cell updates/sec

Title: US-09-091-605-2
Perfect score: 93
Sequence: 1 CATGCTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	93	100.0	359	13	BI715164	BI715164 ic29g03.y
	2	93	100.0	382	13	BM313323	BM313323 ig82f07.y
	3	93	100.0	389	14	BQ632756	BQ632756 il27e09.y
	4	93	100.0	394	12	BG654963	BG654963 ib44a12.y
	5	93	100.0	419	14	BM836042	BM836042 K-EST0111
	6	93	100.0	427	13	BI466966	BI466966 ic17d08.y
	7	93	100.0	443	14	BQ271272	BQ271272 ik11g04.y
	8	93	100.0	451	13	BM503895	BM503895 ig97b05.y
c	9	93	100.0	451	14	BQ776591	BQ776591 il34g04.x
	10	93	100.0	458	12	BG656237	BG656237 ib38g04.y
	11	93	100.0	459	14	BQ286311	BQ286311 ik28e04.y
	12	93	100.0	461	13	BM312520	BM312520 ig75e12.y
	13	93	100.0	463	14	BQ271456	BQ271456 ik14b08.y
	14	93	100.0	466	14	BQ416911	BQ416911 ik39c06.y
	15	93	100.0	468	14	BQ271361	BQ271361 ik12h07.y
	16	93	100.0	471	12	BG655087	BG655087 ib45h03.y
	17	93	100.0	471	14	BQ269449	BQ269449 ik25g03.y
	18	93	100.0	472	14	BQ777016	BQ777016 il42c09.y
	19	93	100.0	472	14	BQ787174	BQ787174 il51c09.y
c	20	93	100.0	474	13	BI467273	BI467273 ic22c02.x
	21	93	100.0	474	14	BQ271282	BQ271282 ik11h07.y
	22	93	100.0	476	14	BQ777432	BQ777432 il48b03.y
	23	93	100.0	478	14	BQ272090	BQ272090 ik18b09.y
	24	93	100.0	479	13	BM312257	BM312257 ig41b09.y
	25	93	100.0	479	14	BQ271314	BQ271314 ik12c06.y
	26	93	100.0	481	13	BM313155	BM313155 ig80c02.y
	27	93	100.0	481	14	BQ776373	BQ776373 il31d12.y
	28	93	100.0	483	13	BI715416	BI715416 ic32f01.y
	29	93	100.0	483	14	BQ269264	BQ269264 ik23d04.y
	30	93	100.0	484	14	BQ271926	BQ271926 ik16b09.y
	31	93	100.0	484	14	BQ417049	BQ417049 ik41b03.y
	32	93	100.0	484	14	BQ776874	BQ776874 il34d03.y
	33	93	100.0	485	14	BQ631915	BQ631915 il24a04.y
	34	93	100.0	485	14	BQ777960	BQ777960 il39c04.y
	35	93	100.0	486	14	BQ778021	BQ778021 il40a04.y
	36	93	100.0	487	13	BM312448	BM312448 ig78f12.x
	37	93	100.0	487	14	BQ269028	BQ269028 ik20b03.y
	38	93	100.0	487	14	BQ269433	BQ269433 ik25e08.y
	39	93	100.0	490	13	BM509838	BM509838 ig94f02.y
	40	93	100.0	492	13	BM312759	BM312759 ig78g03.y
	41	93	100.0	494	13	BI791789	BI791789 ie04e11.y
	42	93	100.0	497	12	BG654541	BG654541 ib41h07.y
c	43	93	100.0	497	13	BI438781	BI438781 ic26f01.x
	44	93	100.0	497	13	BI792098	BI792098 ie06f11.y
	45	93	100.0	501	12	BE970164	BE970164 601680194

ALIGNMENTS

RESULT 1
BI715164
LOCUS BI715164 359 bp mRNA linear EST 19-SEP-2001
DEFINITION ic29g03.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BI715164
VERSION BI715164.1 GI:15690859
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka;I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 277.

```
FEATURES             Location/Qualifiers
     source          1. .359
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      103 a     74 c     92 g     90 t
ORIGIN
```

```
  Query Match             100.0%;  Score 93;  DB 13;  Length 359;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 185

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     186 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 218
```

```
RESULT 2
BM313323
LOCUS       BM313323                382 bp    mRNA    linear   EST 03-JAN-2002
DEFINITION  ig82f07.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN
            P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION   BM313323
VERSION     BM313323.1  GI:18047668
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ig82f07.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
```

```
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      110 a     79 c     98 g     95 t
ORIGIN
```

```
  Query Match             100.0%;  Score 93;  DB 13;  Length 382;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 194

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     195 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 227
```

```
RESULT 3
BQ632756
LOCUS       BQ632756                389 bp    mRNA    linear   EST 02-JUL-2002
DEFINITION  il27e09.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6031216 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ632756
VERSION     BQ632756.1  GI:21684274
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: il27e09.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 361.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /clone="IMAGE:6031216"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      116 a     76 c     94 g    103 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 14;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 131

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     132 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 164


RESULT 4
BG654963
LOCUS       BG654963                394 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  ib44a12.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN
            P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION   BG654963
VERSION     BG654963.1  GI:13792372
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 394)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 350.
FEATURES             Location/Qualifiers
     source          1. .394
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      119 a     74 c     90 g    110 t      1 others
ORIGIN

  Query Match             100.0%;  Score 93;  DB 12;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 117

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     118 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 150


RESULT 5
BM836042
LOCUS       BM836042                419 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0111434 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-18-G08
            5', mRNA sequence.
ACCESSION   BM836042
VERSION     BM836042.1  GI:19192451
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 419)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 18  row: G  column: 08
            High quality sequence stop: 419.
FEATURES             Location/Qualifiers
     source          1. .419
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484s1-18-G08"
                     /clone_lib="S5SNU484s1"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
```

After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promotor as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original liberary and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformaion of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

BASE COUNT 122 a 82 c 94 g 121 t
ORIGIN

Query Match 100.0%; Score 93; DB 14; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 89

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      90 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 122
```

RESULT 6
BI466966
LOCUS BI466966 427 bp mRNA linear EST 22-AUG-2001
DEFINITION ic17d08.y3 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BI466966
VERSION BI466966.1 GI:15279844
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 427)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Sceance,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: ic17d08.x3
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Trace considered overall poor quality
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

BASE COUNT 123 a 94 c 111 g 99 t
ORIGIN

Query Match 100.0%; Score 93; DB 13; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     290 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 349

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     350 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 382
```

RESULT 7
BQ271272
LOCUS BQ271272 443 bp mRNA linear EST 07-MAY-2002
DEFINITION ik11g04.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5780910 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BQ271272
VERSION BQ271272.1 GI:20496338
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Sceance,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco.

FEATURES Location/Qualifiers
source 1. .443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE: 5780910"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University

```
                    School of Medicine, Box 8127, 660 South Euclid Ave., St.
                    Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                    314-362-1916, Fax: 314-747-2692."
BASE COUNT      132 a     99 c    110 g    101 t      1 others
ORIGIN

  Query Match             100.0%;  Score 93;  DB 14;  Length 443;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 394

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    395 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 427


RESULT 8
BM503895
LOCUS       BM503895                451 bp    mRNA    linear   EST 14-FEB-2002
DEFINITION  ig97b05.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN
            P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION   BM503895
VERSION     BM503895.1  GI:18666121
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ig97b05.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 440.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      135 a     86 c    106 g    123 t      1 others
ORIGIN

  Query Match             100.0%;  Score 93;  DB 13;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 133

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    134 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 166


RESULT 9
BQ776591/c
LOCUS       BQ776591                451 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  il34g04.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6032047 3'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ776591
VERSION     BQ776591.1  GI:21985063
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6032047"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      132 a     96 c     82 g    141 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 14;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     451 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 392

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     391 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 359


RESULT 10
BG656237
LOCUS       BG656237                458 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  ib38g04.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN
            P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION   BG656237
VERSION     BG656237.1  GI:13793646
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 452.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      134 a    106 c    111 g    107 t
ORIGIN

  Query Match           100.0%;  Score 93;  DB 12;  Length 458;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     348 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 407

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     408 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 440


RESULT 11
BQ286311
LOCUS       BQ286311                459 bp    mRNA    linear   EST 14-MAY-2002
DEFINITION  ik28e04.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5782351 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ286311
VERSION     BQ286311.1  GI:20655687
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 459)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 440.
FEATURES             Location/Qualifiers
     source          1. .459
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5782351"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      135 a    105 c    109 g    110 t
ORIGIN

  Query Match           100.0%;  Score 93;  DB 14;  Length 459;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 425

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     426 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 458


RESULT 12
```

```
BM312520
LOCUS       BM312520                 461 bp    mRNA    linear   EST 03-JAN-2002
DEFINITION  ig75e12.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN
            P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION   BM312520
VERSION     BM312520.1  GI:18046865
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 401.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      128 a    102 c    118 g    113 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 13;  Length 461;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 324

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     325 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 357


RESULT 13
BQ271456
LOCUS       BQ271456                 463 bp    mRNA    linear   EST 07-MAY-2002
DEFINITION  ik14b08.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5780703 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ271456
VERSION     BQ271456.1  GI:20496522
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 463)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ik14b08.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .463
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE: 5780703"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      135 a    107 c    115 g    106 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 14;  Length 463;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 420

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     421 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 453


RESULT 14
BQ416911
LOCUS       BQ416911                 466 bp    mRNA    linear   EST 23-MAY-2002
DEFINITION  ik39c06.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5783410 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ416911
VERSION     BQ416911.1  GI:21122112
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 466)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ik39c06.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .466
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE: 5783410"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      139 a    108 c    113 g    106 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 14;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 429

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     430 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 462


RESULT 15
BQ271361
LOCUS       BQ271361                468 bp    mRNA    linear   EST 07-MAY-2002
DEFINITION  ik12h07.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5780965 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ271361
VERSION     BQ271361.1  GI:20496427
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ik12h07.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE: 5780965"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      140 a    108 c    114 g    106 t
ORIGIN

  Query Match             100.0%;  Score 93;  DB 14;  Length 468;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 429

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     430 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 462


Search completed: February 14, 2003, 09:05:34
Job time : 1189 secs
```

us-09-091-605-4.rge

Thu Feb 27 13:12:08 2003

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:50:34 ; Search time 621.5 Seconds
(without alignments)
4354.885 Million cell updates/sec

Title: US-09-091-605-4
Perfect score: 93
Sequence: 1 CATGTTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91.4	98.3	396	6	AX147675	AX147675 Sequence
2	91.4	98.3	528	6	AR030615	AR030615 Sequence
3	91.4	98.3	528	6	AR168153	AR168153 Sequence
4	91.4	98.3	528	6	E05860	E05860 DNA encodin
5	91.4	98.3	559	4	AF529185	AF529185 Ovis arie
6	91.4	98.3	955	6	AR108107	AR108107 Sequence
7	91.4	98.3	955	6	AR108109	AR108109 Sequence
8	91.4	98.3	1062	9	HUMGLUC	J04040 Human gluca
9	91.4	98.3	1108	4	BOVGG	K00107 Bovine panc
10	91.4	98.3	1154	9	BC005278	BC005278 Homo sapi
11	91.4	98.3	2356	6	AR108119	AR108119 Sequence
12	91.4	98.3	6455	9	HSGLUC	V01515 Human gene
13	91.4	98.3	10050	9	HSGLUCG2	X03991 Human gluca
c 14	91.4	98.3	163681	9	AC007750	AC007750 Homo sapi
15	89.8	96.6	528	6	AR030614	AR030614 Sequence
16	89.8	96.6	528	6	AR168152	AR168152 Sequence
17	89.8	96.6	1123	4	AF308439	AF308439 Canis fam
18	87.2	93.8	1053	10	GPIGG	D00014 Cavia porce
19	85	91.4	1104	10	OCOGLU	M57688 Octodon deg
20	81.8	88.0	277	10	RATGLU4	K02811 Rat glucago
21	81.8	88.0	895	6	AR108106	AR108106 Sequence
22	81.8	88.0	1034	6	A31421	A31421 H.sapiens m
23	81.8	88.0	1102	10	BC012975	BC012975 Mus muscu
c 24	81.8	88.0	182906	2	AC111919	AC111919 Rattus no
c 25	81.8	88.0	210317	2	AC115074	AC115074 Mus muscu
26	81.8	88.0	219014	2	AC024141	AC024141 Mus muscu
27	80.2	86.2	668	10	MMPPROGLG	Z46845 M.musculus
28	80.2	86.2	1116	10	AF276754	AF276754 Mus muscu
29	79.2	85.2	144	6	AR030619	AR030619 Sequence
30	79.2	85.2	144	6	AR168157	AR168157 Sequence
31	77.4	83.2	106	11	G42680	G42680 GCG Bovine
32	77	82.8	1118	10	HAMGG	J00059 Syrian hams
33	70.6	75.9	1186	5	S78477	S78477 proglucagon
34	70.6	75.9	1576	5	GGPPG	Y07539 Chicken mRN
35	70.6	75.9	1576	6	E03593	E03593 DNA encodin
36	64.2	69.0	1050	5	HSU77612	U77612 Heloderma s
37	64.2	69.0	1196	5	HSU77611	U77611 Heloderma s
38	61.6	66.2	315	6	I07551	I07551 Sequence 40
39	49.8	53.5	78	6	AR077755	AR077755 Sequence
40	48.4	52.0	1408	5	AF004432	AF004432 Xenopus l
41	46	49.5	1290	5	AF004433	AF004433 Xenopus l
42	44.4	47.7	1300	5	AF324209	AF324209 Hoplobatr
43	44	47.3	493	5	LAGL01	V00632 Lophius ame
44	39.8	42.8	383	5	OMU19913	U19913 Oncorhynchu
45	39.8	42.8	383	5	OTU19920	U19920 Oncorhynchu

ALIGNMENTS

RESULT 1
AX147675
LOCUS AX147675 396 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 2 from Patent WO0136643.
ACCESSION AX147675
VERSION AX147675.1 GI:14346730
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 396)
AUTHORS Treco,D.A., Concino,M.F. and Duguay,S.J.
TITLE Nucleic acid construct for optimized production of products
JOURNAL Patent: WO 0136643-A 2 25-MAY-2001;
TRANSKARYOTIC THERAPIES, INC. (US)

```
FEATURES
     source          Location/Qualifiers
                     1. .396
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthesized nucleotide"
BASE COUNT         81 a    111 c    121 g     83 t
ORIGIN

  Query Match                    98.3%;  Score 91.4;  DB 6;  Length 396;
  Best Local Similarity   98.9%;  Pred. No. 5.2e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     286 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 345

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     346 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 378


RESULT 2
AR030615
LOCUS       AR030615                    528 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 8 from patent US 5861284.
ACCESSION   AR030615
VERSION     AR030615.1  GI:5943829
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,O., Kuriyama,M., Koyama,N. and Fukuda,T.
  TITLE     Method for producing a biologically active recombinant
            cysteine-free parathyroid hormone (1-34)
  JOURNAL   Patent: US 5861284-A 8 19-JAN-1999;
FEATURES
     source          Location/Qualifiers
                     1. .528
                     /organism="unknown"
BASE COUNT         142 a    121 c    143 g    122 t
ORIGIN

  Query Match                    98.3%;  Score 91.4;  DB 6;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 5.2e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 3
AR168153
LOCUS       AR168153                    528 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 8 from patent US 6287806.
ACCESSION   AR168153
VERSION     AR168153.1  GI:17903977
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,O., Kuriyama,M., Koyama,N. and Fukuda,T.
  TITLE     Method for producing a biologically active recombinant
            cysteine-free peptide
  JOURNAL   Patent: US 6287806-A 8 11-SEP-2001;
FEATURES
     source          Location/Qualifiers
                     1. .528
                     /organism="unknown"
BASE COUNT         142 a    121 c    143 g    122 t
ORIGIN

  Query Match                    98.3%;  Score 91.4;  DB 6;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 5.2e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 4
E05860
LOCUS       E05860                      528 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  DNA encoding fusion protein comprising glucogon-like peptide 1 and
            fibroblast growth factor.
ACCESSION   E05860
VERSION     E05860.1  GI:2174047
KEYWORDS    JP 1993304976-A/3.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,T., Koyama,N., Kuriyama,M. and Fukuda,T.
  TITLE     PRODUCTION OF PEPTIDE
  JOURNAL   Patent: JP 1993304976-A 3 19-NOV-1993;
COMMENT     TAKEDA CHEM IND LTD
            OS   Artificial gene
            OC   Artificial sequence; Genes.
            OS   Homo sapiens (human)
            PN   JP 1993304976-A/3
            PD   19-NOV-1993
            PF   18-FEB-1992 JP 1992030635
            PR   19-FEB-1991 JP 91P    24841, 18-OCT-1991 JP 91P   271438, PR
            24-OCT-1991 JP 91P   277724
            PI   NISHIMURA TADASHI, KOYAMA NOBUYUKI, KURIYAMA MASATO, PI
            FUKUDA TSUNEHIKO
            PC   C12P21/00,C07K15/00,C12N15/00//C12N15/18,C12N15/62,C12P21/02,
            PC   (C12P21/00,
            PC   C12R1:19),(C12P21/02,C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key                 Location/Qualifiers
            FH
            FT   CDS                 1. .528
            FT                       /product='fusion protein comprising internal
            FT                       fragment of
            FT                       glucogon-like peptide 1 and partial fragment
            FT                       of fibroblast
            FT   mutation            growth factor'
            FT                       replace(96,'T')
            FT                       /product='fusion comprising internal fragment
            FT                       of
            FT                       glucogon-like peptide 1 and partial fragment
            FT                       of fibroblast
            FT   mat_peptide         growth factor'
            FT                       1. .96
            FT                       /product='internal fragment of glucogon-like
            FT   mutation            peptide 1'
            FT                       replace(96,'T')
            FT                       /product='internal fragment of glucogon-like
            FT   mat_peptide         peptide 1'
            FT                       97. .528
                                     /product='partial fragment of fibroblast
                                growth factor'.
FEATURES                Location/Qualifiers
```

```
     source          1. .528
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT      142 a    122 c    143 g    121 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 6;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 5.2e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 5
AF529185
LOCUS       AF529185                559 bp    mRNA    linear   MAM 13-AUG-2002
DEFINITION  Ovis aries preproglucagon mRNA, partial cds.
ACCESSION   AF529185
VERSION     AF529185.1  GI:22212831
KEYWORDS    .
SOURCE      sheep.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Limesand,S.W. and Hay,W.W. Jr.
  TITLE     Characterization of the endocrine pancreas in an ovine placental
            insufficiency IUGR fetus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 559)
  AUTHORS   Limesand,S.W. and Hay,W.W. Jr.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2002) Pediatrics, University of Colorado Health
            Sciences Center, 4200 E 9th Ave, Denver, CO 80262, USA
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
                     /tissue_type="pancreas"
                     /dev_stage="fetal"
     5'UTR           1. .30
     CDS             31. .>559
                     /codon_start=1
                     /product="preproglucagon"
                     /protein_id="AAM94409.1"
                     /db_xref="GI:22212832"
                     /translation="MKSLYFVAGLLVMLAQGSWQHSLQNTEEKSSSFPAPQTDPLGDP
                     DQISEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVNIVEELRRRHADGSFSDEMNTVLD
                     SLATRDFINWLLQTKI"
BASE COUNT      160 a    137 c    138 g    124 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 4;  Length 559;
  Best Local Similarity   98.9%;  Pred. No. 5.3e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    322 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 381

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    382 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 414


RESULT 6
AR108107
LOCUS       AR108107                955 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 57 from patent US 6110707.
ACCESSION   AR108107
VERSION     AR108107.1  GI:12823594
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 955)
  AUTHORS   Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
            Quaade,C., Kruse,F. and McGarry,D.
  TITLE     Recombinant expression of proteins from secretory cell lines
  JOURNAL   Patent: US 6110707-A 57 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .955
                     /organism="unknown"
BASE COUNT      301 a    181 c    203 g    270 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 6;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 5.4e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 7
AR108109
LOCUS       AR108109                955 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 60 from patent US 6110707.
ACCESSION   AR108109
VERSION     AR108109.1  GI:12823596
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 955)
  AUTHORS   Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
            Quaade,C., Kruse,F. and McGarry,D.
  TITLE     Recombinant expression of proteins from secretory cell lines
  JOURNAL   Patent: US 6110707-A 60 29-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .955
                     /organism="unknown"
BASE COUNT      302 a    180 c    202 g    271 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 6;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 5.4e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 8
HUMGLUC
LOCUS       HUMGLUC                 1062 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human glucagon mRNA, complete cds.
ACCESSION   J04040
```

```
VERSION     J04040.1  GI:183269
KEYWORDS    glicentin; glucagon.
SOURCE      Human (neonate) brainstem, cDNA to mRNA, clones BS[3,A,B].
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1062)
  AUTHORS   Drucker,D.J. and Asa,S.
  TITLE     Glucagon gene expression in vertebrate brain
  JOURNAL   J. Biol. Chem. 263 (27), 13475-13478 (1988)
  MEDLINE   88330860
   PUBMED   2901414
FEATURES             Location/Qualifiers
     source          1. .1062
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="2q36-q37"
     gene            1. .1062
                     /gene="GCG"
     mRNA            <1. .1062
                     /gene="GCG"
                     /product="GCG mRNA"
     CDS             38. .580
                     /gene="GCG"
                     /note="preproglucagon"
                     /codon_start=1
                     /protein_id="AAA52567.1"
                     /db_xref="GI:183270"
                     /db_xref="GDB:G00-119-265"
                     /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                     DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                     NLAARDFINWLIQTKITDRK"
     sig_peptide     38. .97
                     /gene="GCG"
                     /note="glucagon signal peptide"
     mat_peptide     98. .304
                     /gene="GCG"
                     /product="glicentin"
     mat_peptide     194. .280
                     /gene="GCG"
                     /product="glucagon"
     mat_peptide     287. .304
                     /gene="GCG"
                     /product="intervening peptide I"
     mat_peptide     311. .421
                     /gene="GCG"
                     /product="glucagon-like peptide I"
     mat_peptide     428. .466
                     /gene="GCG"
                     /product="intervening peptide II"
     mat_peptide     473. .571
                     /gene="GCG"
                     /product="glucagon-like peptide II"
BASE COUNT      340 a    200 c    215 g    307 t
ORIGIN      81 bp upstream of RsaI site; chromosome 2q36-q37.

  Query Match             98.3%;  Score 91.4;  DB 9;  Length 1062;
  Best Local Similarity   98.9%;  Pred. No. 5.4e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     329 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 388

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     389 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 421


RESULT 9
BOVGG
LOCUS       BOVGG                   1108 bp    mRNA    linear   MAM 27-APR-1993
DEFINITION  Bovine pancreas preproglucagon mRNA.
ACCESSION   K00107
VERSION     K00107.1  GI:163081
KEYWORDS    glicentin; glucagon; hormone.
SOURCE      Bovine cdna to pancreatic mrna.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 1108)
  AUTHORS   Lopez,L.C., Frazier,M.L., Su,C.J., Kumar,A. and Saunders,G.F.
  TITLE     Mammalian pancreatic preproglucagon contains three glucagon-related
            peptides
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 80 (18), 5485-5489 (1983)
  MEDLINE   83299996
   PUBMED   6577439
FEATURES             Location/Qualifiers
     source          1. .1108
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
     CDS             91. .633
                     /note="preproglucagon"
                     /codon_start=1
                     /protein_id="AAA30538.1"
                     /db_xref="GI:163082"
                     /translation="MKSLYFVAGLFVMLVQGSWQRSLQNTEEKSSSFPAPQTDPLGDP
                     DQINEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNKNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVNIVEELRRRHADGSFSDEMNTVLD
                     SLATRDFINWLLQTKITDRK"
BASE COUNT      356 a    232 c    213 g    307 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 4;  Length 1108;
  Best Local Similarity   98.9%;  Pred. No. 5.4e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     382 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 441

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     442 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 474


RESULT 10
BC005278
LOCUS       BC005278                1154 bp    mRNA    linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, glucagon, clone MGC:12325 IMAGE:3950435, mRNA,
            complete cds.
ACCESSION   BC005278
VERSION     BC005278.1  GI:13528971
KEYWORDS    MGC.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1154)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
```

```
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 16 Row: k Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4503944.
FEATURES              Location/Qualifiers
     source           1. .1154
                      /organism="Homo sapiens"
                      /db_xref="LocusID:2641"
                      /db_xref="taxon:9606"
                      /clone="MGC:12325 IMAGE:3950435"
                      /tissue_type="Pancreas"
                      /clone_lib="NIH_MGC_78"
                      /lab_host="DH10B"
                      /note="Vector: pDNR-LIB"
     CDS              100. .642
                      /codon_start=1
                      /product="glucagon"
                      /protein_id="AAH05278.1"
                      /db_xref="GI:13528972"
                      /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                      DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGT
                      FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                      NLAARDFINWLIQTKITDRK"
BASE COUNT      391 a    218 c    226 g    319 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 9;  Length 1154;
  Best Local Similarity   98.9%;  Pred. No. 5.4e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     391 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 450

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     451 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 483


RESULT 11
AR108119
LOCUS       AR108119                2356 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 72 from patent US 6110707.
ACCESSION   AR108119
VERSION     AR108119.1  GI:12823606
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2356)
  AUTHORS   Newgard,C.B., Halban,P., Normington,K.D., Clark,S.A., Thigpen,A.E.,
            Quaade,C., Kruse,F. and McGarry,D.
  TITLE     Recombinant expression of proteins from secretory cell lines
  JOURNAL   Patent: US 6110707-A 72 29-AUG-2000;
FEATURES              Location/Qualifiers
     source           1. .2356
                      /organism="unknown"
BASE COUNT      614 a    600 c    581 g    561 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 6;  Length 2356;
  Best Local Similarity   98.9%;  Pred. No. 5.5e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1704 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 1763

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    1764 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 1796


RESULT 12
HSGLUC
LOCUS       HSGLUC                  6455 bp    DNA     linear   PRI 09-FEB-1999
DEFINITION  Human gene encoding preproglucagon. Glucagon is a 29-amino acid
            pancreatic hormone which counteracts the blood glucose-lowering
            action of insulin by stimulating hepatic glycogenolysis and
            gluconeogenesis. Also included in the proglucagon sequence are two
            regions (GLP-1 and GLP-2) which are homologous to glucagon itself
            but not identical.
ACCESSION   V01515
VERSION     V01515.1  GI:31777
KEYWORDS    glucagon; signal peptide.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6455)
  AUTHORS   Bell,G.I., Sanchez-Pescador,R., Laybourn,P.J. and Najarian,R.C.
  TITLE     Exon duplication and divergence in the human preproglucagon gene
  JOURNAL   Nature 304 (5924), 368-371 (1983)
  MEDLINE   83271477
   PUBMED   6877358
COMMENT     Data kindly reviewed (20-SEP-1983) by G.I. Bell.
FEATURES              Location/Qualifiers
     source           1. .6455
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
     CDS              join(168. .259,1832. .1993,3661. .3798,5168. .5315)
                      /codon_start=1
                      /product="precursor"
                      /protein_id="CAA24759.1"
                      /db_xref="GI:31778"
                      /db_xref="SWISS-PROT:P01275"
                      /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                      DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNNIAKRHDEFERHAEGT
                      FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                      NLAARDFINWLIQTKITDR"
     sig_peptide      168. .227
     mat_peptide      1896. .1982
                      /product="glucagon"
     mat_peptide      3680. .3790
                      /product="GLP-1"
     mat_peptide      5211. .5312
                      /product="GLP-2"
     intron           260. .1831
                      /number=1
     intron           1994. .3660
                      /number=2
     intron           3799. .5167
                      /number=3
BASE COUNT     2167 a   1107 c   1121 g   2060 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 9;  Length 6455;
  Best Local Similarity   98.9%;  Pred. No. 5.7e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3698 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 3757

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    3758 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 3790


RESULT 13
HSGLUCG2
```

```
LOCUS       HSGLUCG2                10050 bp    DNA     linear   PRI 20-APR-1995
DEFINITION  Human glucagon gene.
ACCESSION   X03991
VERSION     X03991.1  GI:31786
KEYWORDS    glucagon; glucagon-like peptide; hormone; preproglucagon;
            preprohormone; signal peptide.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10050)
  AUTHORS   White,J.W. and Saunders,G.F.
  TITLE     Structure of the human glucagon gene
  JOURNAL   Nucleic Acids Res. 14 (12), 4719-4730 (1986)
  MEDLINE   86259053
   PUBMED   3725587
COMMENT     See also <HSGLUC; V01515> which is derived from the same library.
            Sequence discrepancies may indicate that the two clones represent
            different alleles or may be due to cloning artefacts. pos. 3516 in
            <X03991> corresponds to pos. 9 in <V01515> (the first eight
            nucleotides of which are probably a cloning artefact).
FEATURES             Location/Qualifiers
     source          1. .10050
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="lambdahGCG1"
                     /clone_lib="fetal liver genomic DNA"
     promoter        complement(535. .538)
                     /note="CAAT-box (complementary strand)"
     promoter        577. .581
                     /note="TATA-box"
     exon            603. .698
                     /note="exon 1; untranslated leader (5' UT-region)"
     intron          699. .3665
                     /note="intron I"
     exon            3666. .3766
                     /note="exon 2"
     CDS             join(3675. .3766,5339. .5500,7177. .7314,8683. .8826,
                     9481. .9487)
                     /note="prepro-glucagon"
                     /codon_start=1
                     /protein_id="CAA27627.1"
                     /db_xref="GI:762941"
                     /db_xref="SWISS-PROT:P01275"
                     /translation="MKSIYFVAGLFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDP
                     DQMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTNRNRNNIAKRHDEFERHAEGT
                     FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVAIVEELGRRHADGSFSDEMNTILD
                     NLAARDFINWLIQTKITDRK"
     sig_peptide     3675. .3734
     intron          3767. .5338
                     /note="intron II"
     exon            5339. .5500
                     /note="exon 3"
     misc_feature    5339. .5396
                     /note="glicentin related pancreatic peptide (AA 12-30)
                     (5339 is 3rd base in codon)"
     misc_feature    5397. .5402
                     /note="spacer (AA 31-32; 2AA)"
     misc_feature    5403. .5489
                     /note="glucagon (AA 33-61; 29AA)"
     misc_feature    5490. .5500
                     /note="spacer (AA 61-65; 4AA) (5500 is 2nd base in codon)"
     intron          5501. .7176
                     /note="intron III"
     exon            7177. .7314
                     /note="exon 4"
     misc_feature    7177. .7195
                     /note="spacer (AA 66-71; 6AA) (7177 is 3rd base in codon)"
     misc_feature    7196. .7306
                     /note="glucagon-like peptide 1 (AA 72-108; 37AA)"
     misc_feature    7307. .7314
                     /note="spacer (AA 109-111; 3AA) (7314 is 2nd base in
                     codon)"
     intron          7315. .8682
                     /note="intron IV"
     exon            8683. .8826
                     /note="exon 5"
     misc_feature    8683. .8826
                     /note="pro-glucagon (AA 112-159) (8683 is 3rd base in
                     codon) (8826 is 2nd base in codon)"
     misc_feature    8683. .8725
                     /note="spacer (AA 112-125; 14AA) (8683 is 3rd base in
                     codon)"
     misc_feature    8726. .8826
                     /note="glucagon-like peptide 2 (AA 126-159) (8826 is 2nd
                     base in codon)"
     intron          8827. .9480
                     /note="intron V"
     exon            9481. .>9487
                     /note="exon 6"
     misc_feature    9481. .9484
                     /note="glucagon-like peptide 2 (AA 159-160; 35AA total)
                     (9481 is 3rd base in codon)"
     misc_feature    9481. .9484
                     /note="pro-glucagon (AA 159-160) (9481 is 3rd base in
                     codon)"
     misc_feature    9950. .9956
                     /note="pot. polyA signal"
     misc_feature    9965. .9974
                     /note="region of pot. polyA sites"
BASE COUNT     3397 a   1698 c   1746 g   3209 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 9;  Length 10050;
  Best Local Similarity   98.9%;  Pred. No. 5.8e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7214 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 7273

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    7274 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 7306


RESULT 14
AC007750/c
LOCUS       AC007750              163681 bp    DNA     linear   PRI 02-OCT-2000
DEFINITION  Homo sapiens BAC clone RP11-576I16 from 2, complete sequence.
ACCESSION   AC007750
VERSION     AC007750.3  GI:6094634
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163681)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 163681)
  AUTHORS   Cotton,M., Maupin,R., Hawkins,M. and Harkins,R.
  TITLE     The sequence of Homo sapiens BAC clone RP11-576I16
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 163681)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 163681)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (22-OCT-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 163681)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Oct 22, 1999 this sequence version replaced gi:5103889.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0576I16
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.
                    .
            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-178A14.  Actual start of
            this clone is at base position 1 of RP11-576I16; actual end is at
            base position 163681 of RP11-576I16.
FEATURES                Location/Qualifiers
     source             1. .163681
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="2"
                        /map="2"
                        /clone="RP11-576I16"
                        /clone_lib="RPCI-11"
     repeat_region      216. .322
                        /rpt_family="MER81"
     repeat_region      1271. .1298
                        /rpt_family="(CAGCT)n"
     repeat_region      1536. .1694
                        /rpt_family="MIR"
     repeat_region      2098. .3518
                        /rpt_family="L2"
     repeat_region      3614. .3902
                        /rpt_family="Alu"
     repeat_region      4090. .4568
                        /rpt_family="MER1_type"
     repeat_region      5128. .5148
                        /rpt_family="AT_rich"
     repeat_region      5263. .5299
                        /rpt_family="(CA)n"
     repeat_region      5417. .5648
                        /rpt_family="L1"
     repeat_region      7093. .7556
                        /rpt_family="L2"
     repeat_region      8702. .9187
                        /rpt_family="MaLR"
     repeat_region      9386. .9586
                        /rpt_family="MIR"
     repeat_region      10580. .10739
                        /rpt_family="L2"
     repeat_region      10792. .11064
                        /rpt_family="Alu"
     repeat_region      11136. .12912
                        /rpt_family="L1"
     misc_feature       12894. .12899
                        /note="match to EST AI077564 (NID:g3411972) oz33g05.x1"
     repeat_region      12982. .13011
                        /rpt_family="(TTG)n"
     repeat_region      13109. .13133
                        /rpt_family="(TTA)n"
     repeat_region      13134. .13419
                        /rpt_family="Alu"
     repeat_region      13833. .14130
                        /rpt_family="Alu"
     repeat_region      14437. .14573
                        /rpt_family="MER1_type"
     repeat_region      14742. .14887
                        /rpt_family="L1"
     repeat_region      15098. .15234
                        /rpt_family="L1"
     misc_feature       15444. .15460
                        /note="match to EST AI922427 (NID:g5658391) wo06b10.x1"
     repeat_region      15449. .15475
                        /rpt_family="(T)n"
     repeat_region      15476. .15695
                        /rpt_family="L1"
     repeat_region      15713. .15993
                        /rpt_family="Alu"
     repeat_region      15994. .16013
                        /rpt_family="(TAAA)n"
     repeat_region      16039. .16131
                        /rpt_family="L1"
     repeat_region      16140. .16281
                        /rpt_family="L1"
     repeat_region      16284. .16327
                        /rpt_family="(TG)n"
     repeat_region      16328. .16394
                        /rpt_family="(CATA)n"
     repeat_region      16396. .16576
                        /rpt_family="L1"
     repeat_region      16625. .16997
                        /rpt_family="L1"
     repeat_region      17020. .17274
                        /rpt_family="L1"
     repeat_region      17270. .17603
                        /rpt_family="L1"
     repeat_region      17644. .18442
                        /rpt_family="L1"
     repeat_region      18444. .20187
                        /rpt_family="L1"
     repeat_region      20260. .20562
                        /rpt_family="MER2_type"
     repeat_region      20563. .20850
                        /rpt_family="Alu"
     repeat_region      20851. .20910
                        /rpt_family="MER2_type"
     repeat_region      21194. .21218
                        /rpt_family="(T)n"
     misc_feature       21204. .21219
                        /note="match to EST AI582416 (NID:g4568313) tr97d11.x1"
     misc_feature       21204. .21219
```

```
                     /note="match to EST AI583407 (NID:g4569304) ts09d01.x1"
     misc_feature    21204. .21219
                     /note="match to EST AI589439 (NID:g4598487) tr76a08.x1"
     misc_feature    21204. .21219
                     /note="match to EST AI865901 (NID:g5530008) wk88b06.x1"
     repeat_region   21618. .21730
                     /rpt_family="MER1_type"
     repeat_region   21800. .21840
                     /rpt_family="AT_rich"
     repeat_region   21991. .22290
                     /rpt_family="Alu"
     misc_feature    21991. .22005
                     /note="match to EST AI682812 (NID:g4892994) wc66f03.x1"
     repeat_region   22464. .22658
                     /rpt_family="MaLR"
     repeat_region   22756. .23030
                     /rpt_family="MaLR"
     repeat_region   23343. .23397
                     /rpt_family="(CA)n"
     repeat_region   23919. .23983
                     /rpt_family="MIR"
     repeat_region   24424. .24581
                     /rpt_family="MER1_type"
     misc_feature    24728. .24983
                     /note="match to EST T10929 (NID:g391083)"
     misc_feature    24730. .25234

  Query Match             98.3%;  Score 91.4;  DB 9;  Length 163681;
  Best Local Similarity   98.9%;  Pred. No. 6.5e-21;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   27495 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 27436

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db   27435 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 27403


RESULT 15
AR030614
LOCUS       AR030614                 528 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 7 from patent US 5861284.
ACCESSION   AR030614
VERSION     AR030614.1  GI:5943828
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Nishimura,O., Kuriyama,M., Koyama,N. and Fukuda,T.
  TITLE     Method for producing a biologically active recombinant
            cysteine-free parathyroid hormone (1-34)
  JOURNAL   Patent: US 5861284-A 7 19-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="unknown"
BASE COUNT      142 a    121 c    144 g    121 t
ORIGIN

  Query Match             96.6%;  Score 89.8;  DB 6;  Length 528;
  Best Local Similarity   97.8%;  Pred. No. 1.9e-20;
  Matches   91;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACGAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
```

Search completed: February 14, 2003, 08:25:24
Job time : 662.5 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 08:25:59 ; Search time 67 Seconds
(without alignments)
706.972 Million cell updates/sec

Title: US-09-091-605-4
Perfect score: 93
Sequence: 1 CATGTTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	91.4	98.3	1174	9	US-09-981-353-169	Sequence 169, App
	2	27.4	29.5	2000	9	US-09-938-842A-4681	Sequence 4681, Ap
	3	27.4	29.5	2002	10	US-09-887-576-22	Sequence 22, Appl
	4	27.2	29.2	490	10	US-09-864-761-345	Sequence 345, App
c	5	27.2	29.2	495	9	US-10-076-622-461	Sequence 461, App
c	6	27.2	29.2	495	10	US-09-604-287A-461	Sequence 461, App
c	7	27.2	29.2	495	12	US-10-007-805-461	Sequence 461, App
	8	27.2	29.2	1856	10	US-09-925-301-395	Sequence 395, App
c	9	27	29.0	29921	9	US-10-083-853-1	Sequence 1, Appli
	10	26.6	28.6	5140	12	US-10-044-090-94	Sequence 94, Appl
c	11	26.4	28.4	2316	10	US-09-801-368-325	Sequence 325, App
	12	26.2	28.2	436	10	US-09-925-297-29	Sequence 29, Appl
c	13	26	28.0	659158	9	US-09-771-208-20	Sequence 20, Appl
c	14	25.8	27.7	2048	12	US-10-044-090-633	Sequence 633, App
c	15	25.8	27.7	2717	12	US-10-044-090-484	Sequence 484, App
c	16	25.8	27.7	23822	10	US-09-964-824A-572	Sequence 572, App
c	17	25.6	27.5	171	10	US-09-864-761-19284	Sequence 19284, A
	18	25.6	27.5	353	10	US-09-867-701-6056	Sequence 6056, Ap
c	19	25.6	27.5	483	10	US-09-864-761-2558	Sequence 2558, Ap
	20	25.6	27.5	1878	10	US-09-925-300-373	Sequence 373, App
	21	25.4	27.3	1194	9	US-09-938-842A-264	Sequence 264, App
	22	25.4	27.3	7970	9	US-09-764-868-1357	Sequence 1357, Ap
	23	25	26.9	1368	10	US-09-942-845-4	Sequence 4, Appli
	24	25	26.9	1368	10	US-09-986-441-4	Sequence 4, Appli
c	25	25	26.9	9543	10	US-09-764-877-2677	Sequence 2677, Ap
c	26	25	26.9	26591	10	US-09-764-877-2678	Sequence 2678, Ap
	27	25	26.9	29729	10	US-09-070-927A-238	Sequence 238, App
c	28	25	26.9	32195	10	US-09-764-869-2017	Sequence 2017, Ap
c	29	25	26.9	32219	10	US-09-764-869-2016	Sequence 2016, Ap
	30	24.8	26.7	197496	9	US-09-877-177-10	Sequence 10, Appl
	31	24.2	26.0	607	10	US-09-770-149-902	Sequence 902, App
	32	24.2	26.0	982	12	US-10-062-254-337	Sequence 337, App
	33	24.2	26.0	2206	12	US-10-062-254-339	Sequence 339, App
	34	24.2	26.0	65608	9	US-09-954-531-180	Sequence 180, App
	35	24.2	26.0	65608	10	US-09-962-436-292	Sequence 292, App
	36	24.2	26.0	65608	10	US-09-962-832-119	Sequence 119, App
c	37	24	25.8	54	9	US-09-328-130-6	Sequence 6, Appli
	38	24	25.8	369	10	US-09-960-352-7411	Sequence 7411, Ap
	39	24	25.8	378	10	US-09-960-352-8269	Sequence 8269, Ap
	40	24	25.8	503	9	US-09-796-692-4557	Sequence 4557, Ap
	41	24	25.8	1242	9	US-09-938-842A-1781	Sequence 1781, Ap
c	42	24	25.8	1337	9	US-10-028-072-459	Sequence 459, App
c	43	24	25.8	1337	9	US-10-121-049-459	Sequence 459, App
c	44	24	25.8	1337	9	US-10-123-904-459	Sequence 459, App
c	45	24	25.8	1337	9	US-10-140-470-459	Sequence 459, App

ALIGNMENTS

RESULT 1
US-09-981-353-169
; Sequence 169, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 169
; LENGTH: 1174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1075717.1
US-09-981-353-169

Query Match 98.3%; Score 91.4; DB 9; Length 1174;
Best Local Similarity 98.9%; Pred. No. 4.1e-24;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     427 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 486

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     487 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 519
```

RESULT 2
US-09-938-842A-4681
; Sequence 4681, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel

```
  ;  APPLICANT:  Wang, Xun
  ;  APPLICANT:  Zhu, Tong
  ;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
  ;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
  ;  FILE REFERENCE: SCRIP1300-3
  ;  CURRENT APPLICATION NUMBER: US/09/938,842A
  ;  CURRENT FILING DATE:  2001-08-24
  ;  PRIOR APPLICATION NUMBER: US 60/227,866
  ;  PRIOR FILING DATE: 2000-08-24
  ;  PRIOR APPLICATION NUMBER: US 60/264,647
  ;  PRIOR FILING DATE: 2001-01-16
  ;  PRIOR APPLICATION NUMBER: US 60/300,111
  ;  PRIOR FILING DATE: 2001-06-22
  ;  NUMBER OF SEQ ID NOS: 5379
  ; SEQ ID NO 4681
  ;    LENGTH: 2000
  ;    TYPE: DNA
  ;    ORGANISM: Arabidopsis thaliana
  US-09-938-842A-4681

    Query Match             29.5%;  Score 27.4;  DB 9;  Length 2000;
    Best Local Similarity   59.7%;  Pred. No. 1.9;
    Matches   46;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

  Qy       2 ATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGG 61
             | ||| ||| ||| |||||| |    | | |   || | ||| ||    |  ||| | ||
  Db    1495 AGGTTAAAGTGACTTTTACCCGATGAGAAGGGAATTGTGTGGCAGATAGAATTGCTAGGG 1554

  Qy      62 AATTCATTGCTTGGCTG 78
             |||   || |||  |||
  Db    1555 AATCTCTTTCTTTTCTG 1571


  RESULT 3
  US-09-887-576-22
  ; Sequence 22, Application US/09887576
  ; Patent No. US20020144047A1
  ; GENERAL INFORMATION:
  ;  APPLICANT: Budworth, P.
  ;  APPLICANT:  Brown, D.
  ;  APPLICANT:  Chang, H.
  ;  APPLICANT:  Zhu, T.
  ;  APPLICANT:  Han, B.
  ;  APPLICANT:  Wang, X.
  ;  APPLICANT:  Cooper, Bret
  ;  TITLE OF INVENTION: Promoters for regulation of plant expression
  ;  FILE REFERENCE: 1360.001US1
  ;  CURRENT APPLICATION NUMBER: US/09/887,576
  ;  CURRENT FILING DATE:  2001-06-25
  ;  PRIOR APPLICATION NUMBER: US 60/213,848
  ;  PRIOR FILING DATE: 2000-06-23
  ;  PRIOR APPLICATION NUMBER: US 60/214,087
  ;  PRIOR FILING DATE: 2000-06-23
  ;  PRIOR APPLICATION NUMBER: US 60/258,692
  ;  PRIOR FILING DATE: 2000-12-29
  ;  NUMBER OF SEQ ID NOS: 875
  ;  SOFTWARE: FastSEQ for Windows Version 4.0
  ; SEQ ID NO 22
  ;    LENGTH: 2002
  ;    TYPE: DNA
  ;    ORGANISM: Arabidopsis thaliana
  US-09-887-576-22

    Query Match             29.5%;  Score 27.4;  DB 10;  Length 2002;
    Best Local Similarity   59.7%;  Pred. No. 1.9;
    Matches   46;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

  Qy       2 ATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGG 61
             | ||| ||| ||| |||||| |    | | |   || | ||| ||    |  ||| | ||
  Db    1491 AGGTTAAAGTGACTTTTACCCGATGAGAAGGGAATTGTGTGGCAGATAGAATTGCTAGGG 1550

  Qy      62 AATTCATTGCTTGGCTG 78
             |||   || |||  |||
  Db    1551 AATCTCTTTCTTTTCTG 1567


  RESULT 4
  US-09-864-761-345/c
  ; Sequence 345, Application US/09864761
  ; Patent No. US20020048763A1
  ; GENERAL INFORMATION:
  ;  APPLICANT: Penn, Sharron G.
  ;  APPLICANT:  Rank, David R.
  ;  APPLICANT:  Hanzel, David K.
  ;  APPLICANT:  Chen, Wensheng
  ;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
  ;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
  ;  FILE REFERENCE: Aeomica-X-1
  ;  CURRENT APPLICATION NUMBER: US/09/864,761
  ;  CURRENT FILING DATE:  2001-05-23
  ;  PRIOR APPLICATION NUMBER: US 60/180,312
  ;  PRIOR FILING DATE: 2000-02-04
  ;  PRIOR APPLICATION NUMBER: US 60/207,456
  ;  PRIOR FILING DATE: 2000-05-26
  ;  PRIOR APPLICATION NUMBER: US 09/632,366
  ;  PRIOR FILING DATE: 2000-08-03
  ;  PRIOR APPLICATION NUMBER: GB 24263.6
  ;  PRIOR FILING DATE: 2000-10-04
  ;  PRIOR APPLICATION NUMBER: US 60/236,359
  ;  PRIOR FILING DATE: 2000-09-27
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00666
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00667
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00664
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00669
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00665
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00668
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00663
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00662
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00661
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: PCT/US01/00670
  ;  PRIOR FILING DATE: 2001-01-30
  ;  PRIOR APPLICATION NUMBER: US 60/234,687
  ;  PRIOR FILING DATE: 2000-09-21
  ;  PRIOR APPLICATION NUMBER: US 09/608,408
  ;  PRIOR FILING DATE: 2000-06-30
  ;  PRIOR APPLICATION NUMBER: US 09/774,203
  ;  PRIOR FILING DATE: 2001-01-29
  ;  NUMBER OF SEQ ID NOS: 49117
  ;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
  ; SEQ ID NO 345
  ;    LENGTH: 490
  ;    TYPE: DNA
  ;    ORGANISM: Homo sapiens
  ;    FEATURE:
  ;    OTHER INFORMATION: MAP TO Z93930.10
  ;    OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =  43
  ;    OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL =  11
  ;    OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =  20
  ;    OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.8
  ;    OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
  ;    OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
  ;    OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
  ;    OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
  ;    OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =  22
  ;    OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =  25
  US-09-864-761-345
```

```
  Query Match             29.2%;  Score 27.2;  DB 10;  Length 490;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     146 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 87

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db      86 CTTTCCCCAGCTGATTAGTGTCTAAGGA 59


RESULT 5
US-10-076-622-461/c
; Sequence 461, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C11
;  CURRENT APPLICATION NUMBER: US/10/076,622
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-076-622-461

  Query Match             29.2%;  Score 27.2;  DB 9;  Length 495;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     158 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 99

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db      98 CTTTCCCCAGCTGATTAGTGTCTAAGGA 71


RESULT 6
US-09-604-287A-461/c
; Sequence 461, Application US/09604287A
; Patent No. US20020064872A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C7
;  CURRENT APPLICATION NUMBER: US/09/604,287A
;  CURRENT FILING DATE:  2000-06-22
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 461
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-604-287A-461

  Query Match             29.2%;  Score 27.2;  DB 10;  Length 495;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     158 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 99

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db      98 CTTTCCCCAGCTGATTAGTGTCTAAGGA 71


RESULT 7
US-10-007-805-461/c
; Sequence 461, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-007-805-461

  Query Match             29.2%;  Score 27.2;  DB 12;  Length 495;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     158 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 99

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db      98 CTTTCCCCAGCTGATTAGTGTCTAAGGA 71


RESULT 8
US-09-925-301-395
; Sequence 395, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 395
;   LENGTH: 1856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1851)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-395

  Query Match             29.2%;  Score 27.2;  DB 10;  Length 1856;
  Best Local Similarity   56.8%;  Pred. No. 2.1;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||    | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db    1107 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 1166

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db    1167 CTTTCCCCAGCTGATTAGTGTCTAAGGA 1194


RESULT 9
US-10-083-853-1/c
; Sequence 1, Application US/10083853
; Patent No. US20020164709A1
; GENERAL INFORMATION:
;  APPLICANT: Affymetrix, Inc
;  APPLICANT:  Shigeta, Ron T
;  APPLICANT:  Siani-Rose, Michael A
;  TITLE OF INVENTION: Nucleic Acid Encoding Growth Factor Protein
;  FILE REFERENCE: 3385.1
;  CURRENT APPLICATION NUMBER: US/10/083,853
;  CURRENT FILING DATE:  2002-02-26
;  PRIOR APPLICATION NUMBER: USSN 60/272,663
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 29921
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-083-853-1

  Query Match             29.0%;  Score 27;  DB 9;  Length 29921;
  Best Local Similarity   60.0%;  Pred. No. 7.7;
  Matches   45;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy       9 AGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCAT 68
           | | | ||||  | ||||| || |||  | | || ||   ||||  |  || |  || ||
Db   24405 AAGAATCTTTTGCTGTGATTTAGGTTAATTTGTGTAAAAACAAGTGGTAAATGTTTTTAT 24346

Qy      69 TGCTTGGCTGGTGAA 83
           |  || | || | ||
Db   24345 TCTTTAGATGTTAAA 24331


RESULT 10
US-10-044-090-94
; Sequence 94, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 94
;   LENGTH: 5140
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1096533.1
US-10-044-090-94

  Query Match             28.6%;  Score 26.6;  DB 12;  Length 5140;
  Best Local Similarity   63.1%;  Pred. No. 5.4;
  Matches   41;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       7 GAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTC 66
           || |  |  ||| |  |||    |||||  ||| ||| || |||    ||||| ||| ||
Db    1380 GAGGACAGATTTTCATGTGTGACAAGTTGGTATCTGGTAGACCAGAAGGCCAAAGAAATC 1439

Qy      67 ATTGC 71
           ||| |
Db    1440 ATTCC 1444


RESULT 11
US-09-801-368-325/c
; Sequence 325, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250A1man, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 325
;   LENGTH: 2316
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-325

  Query Match             28.4%;  Score 26.4;  DB 10;  Length 2316;
  Best Local Similarity   57.1%;  Pred. No. 4.6;
  Matches   48;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||  |   |  |   | || ||  |   ||| ||   |||||| | | ||  |
Db    1853 CATGTTACATTTAAAGTGGATATTGCTGAGACAGCTTGTTGAAAAGGCCCACCAGCGTAA 1794

Qy      61 GAATTCATTGCTTGGCTGGTGAAA 84
           | |||||| | |    ||||||||
Db    1793 GTATTCATAGTTGTTGTGGTGAAA 1770


RESULT 12
US-09-925-297-29
; Sequence 29, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA105
;  CURRENT APPLICATION NUMBER: US/09/925,297
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05989
```

```
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 928
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (64)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (372)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (410)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-29

  Query Match             28.2%;  Score 26.2;  DB 10;  Length 436;
  Best Local Similarity   80.6%;  Pred. No. 2.8;
  Matches   54;  Conservative    0;  Mismatches   10;  Indels    3;  Gaps    2;

Qy      10 GGGACCTTTACCAGTGATGTAAGTT--CTTATTTGGAAGGCCAAGCTG-CCAAGGAATTC 66
           ||  | ||||||||||| ||||||||   |  || |||||| ||||||| |||||| ||||
Db     370 GGNCCTTTTACCAGTGGTGTAAGTTTCTTATTTGGGAAGGNCAAGCTGCCCAAGGGATTC 429

Qy      67 ATTGCTT 73
           |||||||
Db     430 ATTGCTT 436


RESULT 13
US-09-771-208-20/c
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
;  APPLICANT: MEDRANO, JUAN
;  APPLICANT:  BRADFORD, ERIC
;  APPLICANT:  HORVAT, SIMON
;  TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
;  FILE REFERENCE: 407T-923710US
;  CURRENT APPLICATION NUMBER: US/09/771,208
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER: US 08/999,477
;  PRIOR FILING DATE: 1997-12-29
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 659158
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (123459)..(123478)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (602466)..(602485)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (546998)..(547017)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (494715)..(494814)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (390986)..(391005)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (346860)..(346823)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (317174)..(317193)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (280353)..(280373)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (271829)..(271848)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (183872)..(183891)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (170625)..(170645)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   LOCATION: (132680)..(132700)
;   OTHER INFORMATION: n is unidentified a, c, g, or t
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20

  Query Match             28.0%;  Score 26;  DB 9;  Length 659158;
  Best Local Similarity   57.3%;  Pred. No. 59;
  Matches   47;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       2 ATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGG 61
           |||||| |  | |||||  || ||    |||| ||   |   |||||  | |  ||||||
Db  298435 ATGTTGTAAAGGCCTTTGGCATTGGAACAAGTGCTATGTGTTAAGGCAAGGAAGCCAAGT 298376

Qy      62 AATTCATTGCTTGGCTGGTGAA 83
             |  |  || |  | ||  ||
Db  298375 GTTAAAGAGCCTTACAGGGCAA 298354


RESULT 14
US-10-044-090-633/c
; Sequence 633, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 633
;   LENGTH: 2048
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1097794.3
;   NAME/KEY: unsure
;   LOCATION: 934, 936
;   OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-633

  Query Match             27.7%;  Score 25.8;  DB 12;  Length 2048;
  Best Local Similarity   58.4%;  Pred. No. 7.3;
  Matches   45;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy      15 CTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTG 74
           ||| |      ||  ||  ||||||  ||||||   |||  |  |   ||| |    ||||
Db     365 CTTGAGGGCGGACCTAGATTCTTAACTGGAAGTTTAAGAAGTGATCCAATCCCAGTCTTG 306

Qy      75 GCTGGTGAAAGGCCGAG 91
             ||   |||||||| ||
Db     305 CCTTCAGAAAGGCCAAG 289


RESULT 15
```

```
US-10-044-090-484/c
; Sequence 484, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 484
;   LENGTH: 2717
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1097750.1
;   NAME/KEY: unsure
;   LOCATION: 1603-1604
;   OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-484

  Query Match             27.7%;  Score 25.8;  DB 12;  Length 2717;
  Best Local Similarity   58.4%;  Pred. No. 8.2;
  Matches   45;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy      15 CTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTG 74
           ||| |     ||  ||  ||||||  ||||||   |||  |  |   ||| |    ||||
Db    1027 CTTGAGGGCGGACCTAGATTCTTAACTGGAAGTTTAAGAAGTGATCCAATCCCAGTCTTG 968

Qy      75 GCTGGTGAAAGGCCGAG 91
             ||   |||||||| ||
Db     967 CCTTCAGAAAGGCCAAG 951


Search completed: February 14, 2003, 10:24:13
Job time : 238 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 13, 2003, 11:03:17 ; Search time 30 Seconds
(without alignments)
26.401 Million cell updates/sec

Title: US-09-091-605-1
Perfect score: 141
Sequence: 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	133	94.3	30	9	US-10-125-255-1	Sequence 1, Appli
2	133	94.3	30	9	US-09-834-229A-5	Sequence 5, Appli
3	133	94.3	30	10	US-09-851-738-4	Sequence 4, Appli
4	133	94.3	30	10	US-09-805-507-4	Sequence 4, Appli
5	133	94.3	30	10	US-09-859-804-4	Sequence 4, Appli
6	133	94.3	30	10	US-09-982-978-4	Sequence 4, Appli
7	133	94.3	30	10	US-09-953-021B-4	Sequence 4, Appli
8	133	94.3	30	12	US-10-072-540A-4	Sequence 4, Appli
9	133	94.3	31	9	US-09-834-229A-1	Sequence 1, Appli
10	133	94.3	31	10	US-09-754-723-1	Sequence 1, Appli
11	133	94.3	31	10	US-09-420-785A-3	Sequence 3, Appli
12	133	94.3	31	10	US-09-876-388-2	Sequence 2, Appli
13	133	94.3	31	10	US-09-876-388-17	Sequence 17, Appl
14	133	94.3	31	10	US-09-876-388-27	Sequence 27, Appl
15	133	94.3	31	10	US-09-876-388-28	Sequence 28, Appl
16	133	94.3	31	10	US-09-851-738-3	Sequence 3, Appli
17	133	94.3	31	10	US-09-805-507-3	Sequence 3, Appli
18	133	94.3	31	10	US-09-859-804-3	Sequence 3, Appli
19	133	94.3	31	10	US-09-982-978-3	Sequence 3, Appli
20	133	94.3	31	10	US-09-953-021B-3	Sequence 3, Appli
21	133	94.3	31	12	US-10-072-540A-1	Sequence 1, Appli
22	133	94.3	36	10	US-09-851-738-2	Sequence 2, Appli
23	133	94.3	36	10	US-09-805-507-2	Sequence 2, Appli
24	133	94.3	36	10	US-09-859-804-2	Sequence 2, Appli
25	133	94.3	36	10	US-09-982-978-2	Sequence 2, Appli
26	133	94.3	36	10	US-09-953-021B-2	Sequence 2, Appli
27	133	94.3	37	10	US-09-420-785A-2	Sequence 2, Appli
28	133	94.3	37	10	US-09-876-388-1	Sequence 1, Appli
29	133	94.3	37	10	US-09-876-388-16	Sequence 16, Appl
30	133	94.3	37	10	US-09-876-388-25	Sequence 25, Appl
31	133	94.3	37	10	US-09-876-388-26	Sequence 26, Appl
32	133	94.3	37	10	US-09-851-738-1	Sequence 1, Appli
33	133	94.3	37	10	US-09-805-507-1	Sequence 1, Appli
34	133	94.3	37	10	US-09-859-804-1	Sequence 1, Appli
35	133	94.3	37	10	US-09-982-978-1	Sequence 1, Appli
36	133	94.3	37	10	US-09-953-021B-1	Sequence 1, Appli
37	132	93.6	31	10	US-09-876-388-29	Sequence 29, Appl
38	132	93.6	31	10	US-09-876-388-30	Sequence 30, Appl
39	132	93.6	31	12	US-10-072-540A-5	Sequence 5, Appli
40	125	88.7	28	10	US-09-851-738-6	Sequence 6, Appli
41	125	88.7	28	10	US-09-805-507-6	Sequence 6, Appli
42	125	88.7	28	10	US-09-859-804-6	Sequence 6, Appli
43	125	88.7	28	10	US-09-982-978-6	Sequence 6, Appli
44	125	88.7	28	10	US-09-953-021B-6	Sequence 6, Appli
45	125	88.7	29	10	US-09-851-738-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-10-125-255-1
; Sequence 1, Application US/10125255
; Patent No. US20020165342A1
; GENERAL INFORMATION:
; APPLICANT: Galloway, John A
; APPLICANT: Hoffmann, James A
; TITLE OF INVENTION: Glucagon-Like Insulinotropic Peptides, Compositions and Method
; FILE REFERENCE: X-9332E
; CURRENT APPLICATION NUMBER: US/10/125,255
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/573,809
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (30)..(30)
; OTHER INFORMATION: The arginine residue at position 30 is modified so as to repla
; OTHER INFORMATION: the terminal carboxyl group with an amine.
US-10-125-255-1

Query Match 94.3%; Score 133; DB 9; Length 30;
Best Local Similarity 86.7%; Pred. No. 5.3e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
| ||||||||||||| |||| |||||||||
Db 1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30

RESULT 2
US-09-834-229A-5
; Sequence 5, Application US/09834229A
; Publication No. US20030022823A1
; GENERAL INFORMATION:
; APPLICANT: Efendic, Suad
; TITLE OF INVENTION: USE OF GLP-1 OR ANALOGS IN TREATMENT OF MYOCARDIAL INFARCTION

```
;  FILE REFERENCE: X-10822A
;  CURRENT APPLICATION NUMBER: US/09/834,229A
;  CURRENT FILING DATE:  2001-04-12
;  PRIOR APPLICATION NUMBER: US 08/915,918
;  PRIOR FILING DATE: 1997-08-21
;  PRIOR APPLICATION NUMBER: US 06/024,980
;  PRIOR FILING DATE: 1996-08-30
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic construct
US-09-834-229A-5

  Query Match             94.3%;  Score 133;  DB 9;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 5.3e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 3
US-09-851-738-4
; Sequence 4, Application US/09851738
; Patent No. US20020055460A1
; GENERAL INFORMATION:
;  APPLICANT: Coolidge, Thomas R.
;  APPLICANT:  Ehlers, Mario R.W.
;  TITLE OF INVENTION: Metabolic Intervention with GLP-1 to Improve the Function of
;  TITLE OF INVENTION:  Ischemic and Reperfused Tissue
;  FILE REFERENCE: P03660US1
;  CURRENT APPLICATION NUMBER: US/09/851,738
;  CURRENT FILING DATE:  2001-05-09
;  PRIOR APPLICATION NUMBER: 09/302,596
;  PRIOR FILING DATE: 1999-04-30
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: mammalian
US-09-851-738-4

  Query Match             94.3%;  Score 133;  DB 10;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 5.3e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 4
US-09-805-507-4
; Sequence 4, Application US/09805507
; Patent No. US20020098195A1
; GENERAL INFORMATION:
;  APPLICANT: COOLIDGE, THOMAS R.
;  APPLICANT:  EHLERS, MARIO
;  TITLE OF INVENTION: TREATMENT OF ACUTE CORONARY SYNDROME WITH GLP-1
;  FILE REFERENCE: 089187/0395
;  CURRENT APPLICATION NUMBER: US/09/805,507
;  CURRENT FILING DATE:  2001-03-14
;  PRIOR APPLICATION NUMBER: 09/859,804
;  PRIOR FILING DATE: 2001-05-18
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Mammalian GLP
;   OTHER INFORMATION: peptide
US-09-805-507-4

  Query Match             94.3%;  Score 133;  DB 10;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 5.3e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 5
US-09-859-804-4
; Sequence 4, Application US/09859804
; Patent No. US20020107206A1
; GENERAL INFORMATION:
;  APPLICANT: COOLIDGE, THOMAS R.
;  APPLICANT:  EHLERS, MARIO
;  TITLE OF INVENTION: TREATMENT OF ACUTE CORONARY SYNDROME WITH GLP-1
;  FILE REFERENCE: 089187/0395
;  CURRENT APPLICATION NUMBER: US/09/859,804
;  CURRENT FILING DATE:  2001-05-18
;  PRIOR APPLICATION NUMBER: 60/205,239
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Mammalian GLP
;   OTHER INFORMATION: peptide
US-09-859-804-4

  Query Match             94.3%;  Score 133;  DB 10;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 5.3e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 6
US-09-982-978-4
; Sequence 4, Application US/09982978
; Patent No. US20020146405A1
; GENERAL INFORMATION:
;  APPLICANT: COOLIDGE, THOMAS R.
;  APPLICANT:  EHLERS, MARIO
;  TITLE OF INVENTION: TREATMENT OF ACUTE CORONARY SYNDROME WITH GLP-1
;  FILE REFERENCE: 089187/0395
;  CURRENT APPLICATION NUMBER: US/09/982,978
;  CURRENT FILING DATE:  2001-10-22
;  PRIOR APPLICATION NUMBER: 09/859,804
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: 60/205,239
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Mammalian GLP
```

```
;   OTHER INFORMATION: peptide
US-09-982-978-4

  Query Match             94.3%;  Score 133;  DB 10;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 5.3e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 7
US-09-953-021B-4
; Sequence 4, Application US/09953021B
; Patent No. US20020147131A1
; GENERAL INFORMATION:
;  APPLICANT: Coolidge, Thomas L.
;  APPLICANT:  Ehlers, Mario R.W.
;  TITLE OF INVENTION: Metabolic Intervention with GLP-1 to Improve the Function of Isch
;  TITLE OF INVENTION:  Reperfused Skeletal Muscle Tissue
;  FILE REFERENCE: P03660US6
;  CURRENT APPLICATION NUMBER: US/09/953,021B
;  CURRENT FILING DATE:  2001-09-11
;  PRIOR APPLICATION NUMBER: 09/302,596
;  PRIOR FILING DATE: 1999-04-30
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-953-021B-4

  Query Match             94.3%;  Score 133;  DB 10;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 5.3e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 8
US-10-072-540A-4
; Sequence 4, Application US/10072540A
; Patent No. US20020123466A1
; GENERAL INFORMATION:
;  APPLICANT: Hoffmann, James
;  TITLE OF INVENTION: GLP-1 FORMULATIONS
;  FILE REFERENCE: X-11368A
;  CURRENT APPLICATION NUMBER: US/10/072,540A
;  CURRENT FILING DATE:  2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/067,600
;  PRIOR FILING DATE: 1997-12-05
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: AMIDATION
US-10-072-540A-4

  Query Match             94.3%;  Score 133;  DB 12;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 5.3e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 9
US-09-834-229A-1
; Sequence 1, Application US/09834229A
; Publication No. US20030022823A1
; GENERAL INFORMATION:
;  APPLICANT: Efendic, Suad
;  TITLE OF INVENTION: USE OF GLP-1 OR ANALOGS IN TREATMENT OF MYOCARDIAL INFARCTION
;  FILE REFERENCE: X-10822A
;  CURRENT APPLICATION NUMBER: US/09/834,229A
;  CURRENT FILING DATE:  2001-04-12
;  PRIOR APPLICATION NUMBER: US 08/915,918
;  PRIOR FILING DATE: 1997-08-21
;  PRIOR APPLICATION NUMBER: US 06/024,980
;  PRIOR FILING DATE: 1996-08-30
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-834-229A-1

  Query Match             94.3%;  Score 133;  DB 9;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 10
US-09-754-723-1
; Sequence 1, Application US/09754723
; Patent No. US20010002394A1
; GENERAL INFORMATION:
;  APPLICANT: EFENDIC, Suad
;  APPLICANT:  GUTNIAK, Mark
;  APPLICANT:  KIRK, Ole
;  TITLE OF INVENTION: Use Of A Peptide
;  FILE REFERENCE: 3745.234-US
;  CURRENT APPLICATION NUMBER: US/09/754,723
;  CURRENT FILING DATE:  2001-06-21
;  PRIOR APPLICATION NUMBER: US 08/842,121
;  PRIOR FILING DATE: 1997-04-23
;  PRIOR APPLICATION NUMBER: US 08/295,913
;  PRIOR FILING DATE: 1994-10-13
;  PRIOR APPLICATION NUMBER: PCT/DK93/00099
;  PRIOR FILING DATE: 1993-03-19
;  PRIOR APPLICATION NUMBER: DK 0363/92
;  PRIOR FILING DATE: 1992-03-19
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Homo Sapien
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(31)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-754-723-1

  Query Match             94.3%;  Score 133;  DB 10;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30
```

```
RESULT 11
US-09-420-785A-3
; Sequence 3, Application US/09420785A
; Patent No. US20010010923A1
; GENERAL INFORMATION:
;  APPLICANT: MORTENSEN, UFFE
;  APPLICANT:  OLESEN, KJELD
;  APPLICANT:  STENNICKE, HENNING
;  APPLICANT:  SORENSEN, STEEN B.
;  APPLICANT:  BREDDAM, KLAUS
;  TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
;  FILE REFERENCE: 089187/0109
;  CURRENT APPLICATION NUMBER: US/09/420,785A
;  CURRENT FILING DATE:  1999-10-19
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (31)
;   OTHER INFORMATION: C-terminal amino acid which serves as a leaving
;   OTHER INFORMATION: group, typically, an uncharged amino acid side
;   OTHER INFORMATION: chain, preferably alanine
US-09-420-785A-3

  Query Match             94.3%;  Score 133;  DB 10;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 12
US-09-876-388-2
; Sequence 2, Application US/09876388
; Patent No. US20020049153A1
; GENERAL INFORMATION:
;  APPLICANT: Bridon, Dominique P.
;  APPLICANT:  L'Archeveque, Benoit
;  APPLICANT:  Ezrin, Alan M.
;  APPLICANT:  Holmes, Darren L.
;  APPLICANT:  Leblanc, Anouk
;  APPLICANT:  St. Pierre, Serge
;  TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
;  FILE REFERENCE: 500862001610
;  CURRENT APPLICATION NUMBER: US/09/876,388
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 09/623,618
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: PCT/US00/13563
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/159,783
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: 60/134,406
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-876-388-2

  Query Match             94.3%;  Score 133;  DB 10;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 13
US-09-876-388-17
; Sequence 17, Application US/09876388
; Patent No. US20020049153A1
; GENERAL INFORMATION:
;  APPLICANT: Bridon, Dominique P.
;  APPLICANT:  L'Archeveque, Benoit
;  APPLICANT:  Ezrin, Alan M.
;  APPLICANT:  Holmes, Darren L.
;  APPLICANT:  Leblanc, Anouk
;  APPLICANT:  St. Pierre, Serge
;  TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
;  FILE REFERENCE: 500862001610
;  CURRENT APPLICATION NUMBER: US/09/876,388
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 09/623,618
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: PCT/US00/13563
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/159,783
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: 60/134,406
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-876-388-17

  Query Match             94.3%;  Score 133;  DB 10;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 14
US-09-876-388-27
; Sequence 27, Application US/09876388
; Patent No. US20020049153A1
; GENERAL INFORMATION:
;  APPLICANT: Bridon, Dominique P.
;  APPLICANT:  L'Archeveque, Benoit
;  APPLICANT:  Ezrin, Alan M.
;  APPLICANT:  Holmes, Darren L.
;  APPLICANT:  Leblanc, Anouk
;  APPLICANT:  St. Pierre, Serge
;  TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
;  FILE REFERENCE: 500862001610
;  CURRENT APPLICATION NUMBER: US/09/876,388
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 09/623,618
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: PCT/US00/13563
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/159,783
;  PRIOR FILING DATE: 1999-10-15
```

```
;  PRIOR APPLICATION NUMBER: 60/134,406
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
;   NAME/KEY: MOD_RES
;   LOCATION: 31
;   OTHER INFORMATION: Xaa represents Lys(E-MPA)-NH2-4TFA and where "E" represents Epsil
US-09-876-388-27

  Query Match             94.3%;  Score 133;  DB 10;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||  ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 15
US-09-876-388-28
; Sequence 28, Application US/09876388.
; Patent No. US20020049153A1
; GENERAL INFORMATION:
;  APPLICANT: Bridon, Dominique P.
;  APPLICANT:  L'Archeveque, Benoit
;  APPLICANT:  Ezrin, Alan M.
;  APPLICANT:  Holmes, Darren L.
;  APPLICANT:  Leblanc, Anouk
;  APPLICANT:  St. Pierre, Serge
;  TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
;  FILE REFERENCE: 500862001610
;  CURRENT APPLICATION NUMBER: US/09/876,388
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 09/623,618
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: PCT/US00/13563
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/159,783
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: 60/134,406
;  PRIOR FILING DATE: 1999-05-17
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
;   NAME/KEY: MOD_RES
;   LOCATION: 31
;   OTHER INFORMATION: Xaa represents Lys(E-AEEA-AEEA-MPA)-NH2-4TFA and where "E" repres
US-09-876-388-28

  Query Match             94.3%;  Score 133;  DB 10;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||  ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


Search completed: February 13, 2003, 11:07:20
Job time : 30 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 08:25:59 ; Search time 67 Seconds
(without alignments)
706.972 Million cell updates/sec

Title: US-09-091-605-2
Perfect score: 93
Sequence: 1 CATGCTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	93	100.0	1174	9	US-09-981-353-169	Sequence 169, App
c	2	27.2	29.2	490	10	US-09-864-761-345	Sequence 345, App
c	3	27.2	29.2	495	9	US-10-076-622-461	Sequence 461, App
c	4	27.2	29.2	495	10	US-09-604-287A-461	Sequence 461, App
c	5	27.2	29.2	495	12	US-10-007-805-461	Sequence 461, App
	6	27.2	29.2	1856	10	US-09-925-301-395	Sequence 395, App
	7	27	29.0	436	10	US-09-925-297-29	Sequence 29, Appl
c	8	27	29.0	29921	9	US-10-083-853-1	Sequence 1, Appli
	9	26.6	28.6	5140	12	US-10-044-090-94	Sequence 94, Appl
	10	25.8	27.7	2000	9	US-09-938-842A-4681	Sequence 4681, Ap
	11	25.8	27.7	2002	10	US-09-887-576-22	Sequence 22, Appl
c	12	25.8	27.7	2048	12	US-10-044-090-633	Sequence 633, App
c	13	25.8	27.7	2717	12	US-10-044-090-484	Sequence 484, App
c	14	25.8	27.7	23822	10	US-09-964-824A-572	Sequence 572, App
c	15	25.6	27.5	54	9	US-09-328-130-6	Sequence 6, Appli
c	16	25.6	27.5	171	10	US-09-864-761-19284	Sequence 19284, A
	17	25.6	27.5	353	10	US-09-867-701-6056	Sequence 6056, Ap
c	18	25.6	27.5	483	10	US-09-864-761-2558	Sequence 2558, Ap
	19	25.6	27.5	1878	10	US-09-925-300-373	Sequence 373, App
	20	25.4	27.3	1194	9	US-09-938-842A-264	Sequence 264, App
	21	25.4	27.3	1614	10	US-09-923-246-7	Sequence 7, Appli
	22	25.4	27.3	1614	10	US-09-825-561A-1	Sequence 1, Appli
	23	25.4	27.3	1707	10	US-09-758-664-3	Sequence 3, Appli
	24	25.4	27.3	1707	10	US-09-825-561A-64	Sequence 64, Appl
	25	25.4	27.3	1740	10	US-09-758-664-1	Sequence 1, Appli
	26	25.4	27.3	2343	10	US-09-965-313-1	Sequence 1, Appli
	27	25.4	27.3	7970	9	US-09-764-868-1357	Sequence 1357, Ap
	28	25	26.9	1368	10	US-09-942-845-4	Sequence 4, Appli
	29	25	26.9	1368	10	US-09-986-441-4	Sequence 4, Appli
c	30	25	26.9	9543	10	US-09-764-877-2677	Sequence 2677, Ap
c	31	25	26.9	26591	10	US-09-764-877-2678	Sequence 2678, Ap
	32	25	26.9	29729	10	US-09-070-927A-238	Sequence 238, App
c	33	24.8	26.7	2316	10	US-09-801-368-325	Sequence 325, App
c	34	24.4	26.2	659158	9	US-09-771-208-20	Sequence 20, Appl
	35	24.2	26.0	607	10	US-09-770-149-902	Sequence 902, App
	36	24.2	26.0	982	12	US-10-062-254-337	Sequence 337, App
	37	24.2	26.0	2206	12	US-10-062-254-339	Sequence 339, App
	38	24.2	26.0	65608	9	US-09-954-531-180	Sequence 180, App
	39	24.2	26.0	65608	10	US-09-962-436-292	Sequence 292, App
	40	24.2	26.0	65608	10	US-09-962-832-119	Sequence 119, App
	41	24	25.8	34	9	US-09-328-130-3	Sequence 3, Appli
	42	24	25.8	54	9	US-09-328-130-5	Sequence 5, Appli
	43	24	25.8	369	10	US-09-960-352-7411	Sequence 7411, Ap
	44	24	25.8	378	10	US-09-960-352-8269	Sequence 8269, Ap
	45	24	25.8	503	9	US-09-796-692-4557	Sequence 4557, Ap

ALIGNMENTS

RESULT 1
US-09-981-353-169
; Sequence 169, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 169
; LENGTH: 1174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1075717.1
US-09-981-353-169

Query Match 100.0%; Score 93; DB 9; Length 1174;
Best Local Similarity 100.0%; Pred. No. 1.1e-24;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     427 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 486

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     487 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 519
```

RESULT 2
US-09-864-761-345/c
; Sequence 345, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

```
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY  .
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 345
;   LENGTH: 490
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO Z93930.10
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =  43
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL =  11
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =  20
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =  22
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =  25
US-09-864-761-345

  Query Match             29.2%;  Score 27.2;  DB 10;  Length 490;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     146 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 87

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
```

```
Db      86 CTTTCCCCAGCTGATTAGTGTCTAAGGA 59


RESULT 3
US-10-076-622-461/c
; Sequence 461, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C11
;  CURRENT APPLICATION NUMBER: US/10/076,622
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-076-622-461

  Query Match             29.2%;  Score 27.2;  DB 9;  Length 495;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     158 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 99

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db      98 CTTTCCCCAGCTGATTAGTGTCTAAGGA 71


RESULT 4
US-09-604-287A-461/c
; Sequence 461, Application US/09604287A
; Patent No. US20020064872A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C7
;  CURRENT APPLICATION NUMBER: US/09/604,287A
;  CURRENT FILING DATE:  2000-06-22
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 461
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-604-287A-461

  Query Match             29.2%;  Score 27.2;  DB 10;  Length 495;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     158 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 99

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db      98 CTTTCCCCAGCTGATTAGTGTCTAAGGA 71
```

```
RESULT 5
US-10-007-805-461/c
; Sequence 461, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-007-805-461

  Query Match             29.2%;  Score 27.2;  DB 12;  Length 495;
  Best Local Similarity   56.8%;  Pred. No. 1.3;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db     158 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 99

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db      98 CTTTCCCCAGCTGATTAGTGTCTAAGGA 71


RESULT 6
US-09-925-301-395
; Sequence 395, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 395
;   LENGTH: 1856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1851)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-395

  Query Match             29.2%;  Score 27.2;  DB 10;  Length 1856;
  Best Local Similarity   56.8%;  Pred. No. 2.2;
  Matches   50;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       6 TGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATT 65
           |||   | ||| |     || |||||    || || ||| | |     |||||| ||| |
Db    1107 TGACATGTCCTCTCTGCTTGGTGTAAACCATTCTTGGGAGGACACTTTTGCCAATGAACT 1166

Qy      66 CATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           | || |   |||| | |  | |  ||||
Db    1167 CTTTCCCCAGCTGATTAGTGTCTAAGGA 1194


RESULT 7
US-09-925-297-29
; Sequence 29, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA105
;  CURRENT APPLICATION NUMBER: US/09/925,297
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05989
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 928
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (64)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (372)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (410)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-29

  Query Match             29.0%;  Score 27;  DB 10;  Length 436;
  Best Local Similarity   57.8%;  Pred. No. 1.5;
  Matches   48;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||| |  | || || || ||||||||    ||   ||| ||||  |    |||| |
Db     222 CATTCACAGGGCACATTCACCAGTGACTACAGCAAGTATCTGGACTCCAGGCGTGCCCAA 281

Qy      61 GAATTCATTGCTTGGCTGGTGAA 83
           || ||  |    ||| || ||||
Db     282 GATTTTGTGCAGTGGTTGATGAA 304


RESULT 8
US-10-083-853-1/c
; Sequence 1, Application US/10083853
; Patent No. US20020164709A1
; GENERAL INFORMATION:
;  APPLICANT: Affymetrix, Inc
;  APPLICANT:  Shigeta, Ron T
;  APPLICANT:  Siani-Rose, Michael A
;  TITLE OF INVENTION: Nucleic Acid Encoding Growth Factor Protein
;  FILE REFERENCE: 3385.1
;  CURRENT APPLICATION NUMBER: US/10/083,853
;  CURRENT FILING DATE:  2002-02-26
;  PRIOR APPLICATION NUMBER: USSN 60/272,663
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 29921
;   TYPE: DNA
```

```
;   ORGANISM: Homo Sapiens
US-10-083-853-1

  Query Match            29.0%;  Score 27;  DB 9;  Length 29921;
  Best Local Similarity  60.0%;  Pred. No. 7.4;
  Matches   45;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy       9 AGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCAT 68
Db   24405 AAGAATCTTTGCTGTGATTTAGGTTAATTTGTGTAAAAACAAGTGGTAAATGTTTTTAT 24346

Qy      69 TGCTTGGCTGGTGAA 83
Db   24345 TCTTTAGATGTTAAA 24331


RESULT 9
US-10-044-090-94
; Sequence 94, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 94
;   LENGTH: 5140
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1096533.1
US-10-044-090-94

  Query Match            28.6%;  Score 26.6;  DB 12;  Length 5140;
  Best Local Similarity  63.1%;  Pred. No. 5.3;
  Matches   41;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       7 GAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTC 66
Db    1380 GAGGACAGATTTTCATGTGTGACAAGTTGGTATCTGGTAGACCAGAAGGCCAAAGAAATC 1439

Qy      67 ATTGC 71
Db    1440 ATTCC 1444


RESULT 10
US-09-938-842A-4681
; Sequence 4681, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4681
;   LENGTH: 2000
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-4681

  Query Match            27.7%;  Score 25.8;  DB 9;  Length 2000;
  Best Local Similarity  58.4%;  Pred. No. 7.2;
  Matches   45;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       2 ATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGG 61
Db    1495 AGGTTAAAGTGACTTTTACCCGATGAGAAGGGAATTGTGTGGCAGATAGAATTGCTAGGG 1554

Qy      62 AATTCATTGCTTGGCTG 78
Db    1555 AATCTCTTTCTTTTCTG 1571


RESULT 11
US-09-887-576-22
; Sequence 22, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, P.
;  APPLICANT:  Brown, D.
;  APPLICANT:  Chang, H.
;  APPLICANT:  Zhu, T.
;  APPLICANT:  Han, B.
;  APPLICANT:  Wang, X.
;  APPLICANT:  Cooper, Bret
;  TITLE OF INVENTION: Promoters for regulation of plant expression
;  FILE REFERENCE: 1360.001US1
;  CURRENT APPLICATION NUMBER: US/09/887,576
;  CURRENT FILING DATE:  2001-06-25
;  PRIOR APPLICATION NUMBER: US 60/213,848
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/214,087
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/258,692
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 875
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 2002
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-887-576-22

  Query Match            27.7%;  Score 25.8;  DB 10;  Length 2002;
  Best Local Similarity  58.4%;  Pred. No. 7.3;
  Matches   45;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       2 ATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGG 61
Db    1491 AGGTTAAAGTGACTTTTACCCGATGAGAAGGGAATTGTGTGGCAGATAGAATTGCTAGGG 1550

Qy      62 AATTCATTGCTTGGCTG 78
Db    1551 AATCTCTTTCTTTTCTG 1567


RESULT 12
US-10-044-090-633/c
; Sequence 633, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 633
```

```
;   LENGTH: 2048
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1097794.3
;   NAME/KEY: unsure
;   LOCATION: 934, 936
;   OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-633

  Query Match            27.7%;  Score 25.8;  DB 12;  Length 2048;
  Best Local Similarity   58.4%;  Pred. No. 7.3;
  Matches   45;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy      15 CTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTG 74
           ||| |     ||  ||  ||||||  ||||||   |||  |  |   ||| |    ||||
Db     365 CTTGAGGGCGGACCTAGATTCTTAACTGGAAGTTTAAGAAGTGATCCAATCCCAGTCTTG 306

Qy      75 GCTGGTGAAAGGCCGAG 91
             ||   |||||||| ||
Db     305 CCTTCAGAAAGGCCAAG 289


RESULT 13
US-10-044-090-484/c
; Sequence 484, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 484
;   LENGTH: 2717
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1097750.1
;   NAME/KEY: unsure
;   LOCATION: 1603-1604
;   OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-484

  Query Match            27.7%;  Score 25.8;  DB 12;  Length 2717;
  Best Local Similarity   58.4%;  Pred. No. 8.2;
  Matches   45;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy      15 CTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTG 74
           ||| |     ||  ||  ||||||  ||||||   |||  |  |   ||| |    ||||
Db    1027 CTTGAGGGCGGACCTAGATTCTTAACTGGAAGTTTAAGAAGTGATCCAATCCCAGTCTTG 968

Qy      75 GCTGGTGAAAGGCCGAG 91
             ||   |||||||| ||
Db     967 CCTTCAGAAAGGCCAAG 951


RESULT 14
US-09-964-824A-572/c
; Sequence 572, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
;  APPLICANT: Horrigan, Stephen
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-73
;  CURRENT APPLICATION NUMBER: US/09/964,824A
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: US/60/236,033
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,032
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,028
;  PRIOR FILING DATE: 2000-09-28
;  NUMBER OF SEQ ID NOS: 583
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 572
;   LENGTH: 23822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-964-824A-572

  Query Match            27.7%;  Score 25.8;  DB 10;  Length 23822;
  Best Local Similarity   58.4%;  Pred. No. 19;
  Matches   45;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy      15 CTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTG 74
           ||| |     ||  ||  ||||||  ||||||   |||  |  |   ||| |    ||||
Db  21710 CTTGAGGGCGGACCTAGATTCTTAACTGGAAGTTTAAGAAGTGATCCAATCCCAGTCTTG 21651

Qy      75 GCTGGTGAAAGGCCGAG 91
             ||   |||||||| ||
Db  21650 CCTTCAGAAAGGCCAAG 21634


RESULT 15
US-09-328-130-6/c
; Sequence 6, Application US/09328130
; Patent No. US20020155597A1
; GENERAL INFORMATION:
;  APPLICANT: Selden, Richard F
;  APPLICANT:  Treco, Douglas
;  APPLICANT:  Heartlein, Michael W.
;  TITLE OF INVENTION: In Vivo Production and Delivery of
;  TITLE OF INVENTION:  Erythropoietin or Insulinotropin for Gene Therapy
;  FILE REFERENCE: 50010/007006
;  CURRENT APPLICATION NUMBER: US/09/328,130
;  CURRENT FILING DATE:  1999-06-08
;  PRIOR APPLICATION NUMBER: US 08/334,455
;  PRIOR FILING DATE: 1994-11-04
;  PRIOR APPLICATION NUMBER: US 07/911,533
;  PRIOR FILING DATE: 1992-07-10
;  PRIOR APPLICATION NUMBER: US 07/787,840
;  PRIOR FILING DATE: 1991-11-05
;  PRIOR APPLICATION NUMBER: US 07/789,188
;  PRIOR FILING DATE: 1991-11-05
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 54
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-328-130-6

  Query Match            27.5%;  Score 25.6;  DB 9;  Length 54;
  Best Local Similarity   87.5%;  Pred. No. 2.1;
  Matches   28;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAG 32
           ||||||||||||||||||||||||||||   || |
Db      34 CATGCTGAAGGGACCTTTACCAGTGTAATATG 3


Search completed: February 14, 2003, 10:21:22
Job time : 77 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 13, 2003, 11:02:01 ; Search time 16 Seconds
(without alignments)
186.260 Million cell updates/sec

Title: US-09-091-605-1
Perfect score: 141
Sequence: 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	133	94.3	158	1	GCPG	glucagon precursor
2	133	94.3	180	1	GCHU	glucagon precursor
3	133	94.3	180	1	GCGP	glucagon precursor
4	133	94.3	180	1	GCRTDU	glucagon precursor
5	133	94.3	180	1	GCRT	glucagon precursor
6	133	94.3	180	1	GCHY	glucagon precursor
7	133	94.3	180	1	GCBO	glucagon precursor
8	133	94.3	180	2	A57294	glucagon precursor
9	121	85.8	151	1	GCCH	glucagon precursor
10	121	85.8	206	2	I51301	proglucagon - chic
11	107	75.9	30	2	B61125	glucagon-like pept
12	107	75.9	30	2	C61125	glucagon-like pept
13	107	75.9	101	1	GCFGB	glucagon precursor
14	104	73.8	122	1	GCAF2	glucagon 2 precurs
15	102	72.3	66	2	I51093	glucagon - chinook
16	102	72.3	178	2	I51058	glucagon I precurs
17	101	71.6	63	1	GCIDC	glucagon precursor
18	101	71.6	72	1	GCGXA	glucagon precursor
19	98	69.5	30	2	S44473	glucagon-like pept
20	97	68.8	60	1	GCONC	glucagon precursor
21	97	68.8	178	2	I51057	glucagon II precur
22	89	63.1	87	1	GCFIS	glucagon precursor
23	84	59.6	29	2	S07211	glucagon - marbled
24	82	58.2	29	1	GCDF	glucagon - smaller
25	81	57.4	29	1	GCOPV	glucagon - North A
26	81	57.4	29	2	A91740	glucagon - turkey
27	81	57.4	29	2	A91741	glucagon - rabbit
28	81	57.4	29	2	A91742	glucagon - Arabian
29	81	57.4	29	2	C39258	glucagon - common
30	81	57.4	69	1	GCDG69	glucagon-69 - dog
31	81	57.4	124	1	GCAF	glucagon 1 precurs
32	80	56.7	29	1	GCEN	glucagon - elephan
33	79	56.0	29	1	GCDK	glucagon - duck
34	79	56.0	29	1	A61583	glucagon - ostrich
35	79	56.0	29	1	GCTTS	glucagon - slider
36	79	56.0	39	1	HWGH4G	exendin-4 - Gila m
37	77	54.6	31	2	S44472	glucagon G2 - Nort
38	77	54.6	39	1	HWGH3Z	exendin-3 - Mexica
39	76	53.9	29	2	C60840	glucagon I - Europ
40	75	53.2	29	1	GCCB	glucagon - Chinchi
41	75	53.2	31	2	S44471	glucagon G1 - Nort
42	74	52.5	29	1	GCFLE	glucagon - Europea
43	74	52.5	29	2	A61135	glucagon - bigeye
44	74	52.5	29	2	S39018	glucagon - bowfin
45	67	47.5	36	1	GCFI	glucagon-36 - spot

ALIGNMENTS

RESULT 1
GCPG
glucagon precursor - pig (fragment)
N;Alternate names: glicentin; oxyntomodulin
N;Contains: glicentin-related peptide; glucagon; glucagon-37 (oxyntomodulin); glucago
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 17-Dec-1982 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C;Accession: A01540; A60312; A91781; B32614; A28064
R;Thim, L.; Moody, A.J.
Regul. Pept. 2, 139-150, 1981
A;Title: The primary structure of porcine glicentin (proglucagon).
A;Reference number: A94233; MUID:81248172; PMID:6894800
A;Accession: A01540
A;Molecule type: protein
A;Residues: 1-69 <TH1>
R;Thim, L.; Moody, A.J.
Regul. Pept. Suppl. 2, S33, 1983
A;Title: Primary structure of a possible porcine proglucagon fragment.
A;Reference number: A60312
A;Accession: A60312
A;Molecule type: protein
A;Residues: 1-30 <TH2>
A;Note: this peptide is co-secreted with glucagon from the pancreas
R;Bromer, W.W.; Sinn, L.G.; Behrens, O.K.
J. Am. Chem. Soc. 79, 2807-2810, 1957
A;Title: The amino acid sequence of glucagon. V. Location of amide groups, acid degra
A;Reference number: A91781
A;Accession: A91781
A;Molecule type: protein
A;Residues: 33-61 <BRO>
R;Orskov, C.; Bersani, M.; Johnsen, A.H.; Hojrup, P.; Holst, J.J.
J. Biol. Chem. 264, 12826-12829, 1989
A;Title: Complete sequences of glucagon-like peptide-1 from human and pig small intes
A;Reference number: A92732; MUID:89327238; PMID:2753890
A;Accession: B32614
A;Molecule type: protein
A;Residues: 78-107 <ORS>
R;Buhl, T.; Thim, L.; Kofod, H.; Orskov, C.; Harling, H.; Holst, J.J.
J. Biol. Chem. 263, 8621-8624, 1988
A;Title: Naturally occurring products of proglucagon 111-160 in the porcine and human
A;Reference number: A28064; MUID:88243712; PMID:3379036
A;Accession: A28064
A;Molecule type: protein
A;Residues: 111-158 <BUH>
C;Comment: X's represent missing amino acids, mostly basic, that are predicted to exi
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; int
F;1-69/Product: glucagon-69 #status experimental <G69>
F;1-30/Region: glicentin-related peptide #status experimental
F;33-69/Product: glucagon-37 #status predicted <G37>
F;33-61/Product: glucagon #status experimental <GCN>
F;78-107/Product: glucagon-like peptide 1 #status experimental <GL1>

```
F;126-158/Product: glucagon-like peptide 2 #status experimental <GL2>
F;107/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             94.3%;  Score 133;  DB 1;  Length 158;
  Best Local Similarity   86.7%;  Pred. No. 5.5e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      78 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 107


RESULT 2
GCHU
glucagon precursor [validated] - human
N;Contains: glicentin; glicentin-related polypeptide (GRPP); glucagon; glucagon-like pep
ke peptide 1 (tGLP1)
C;Species: Homo sapiens (man)
C;Date: 24-Apr-1984 #sequence_revision 31-Mar-1993 #text_change 08-Dec-2000
C;Accession: A24377; A44197; A30875; A32614; A01541; S23309
R;White, J.W.; Saunders, G.F.
Nucleic Acids Res. 14, 4719-4730, 1986
A;Title: Structure of the human glucagon gene.
A;Reference number: A24377; MUID:86259053; PMID:3725587
A;Accession: A24377
A;Molecule type: DNA
A;Residues: 1-180 <WHI>
A;Cross-references: GB:X03991
R;Bell, G.I.; Sanchez-Pescador, R.; Laybourn, P.J.; Najarian, R.C.
Nature 304, 368-371, 1983
A;Title: Exon duplication and divergence in the human preproglucagon gene.
A;Reference number: A44197; MUID:83271477; PMID:6877358
A;Accession: A44197
A;Molecule type: DNA
A;Residues: 1-179 <BEL>
A;Cross-references: GB:V01515; NID:g31777; PIDN:CAA24759.1; PID:g31778
R;Drucker, D.J.; Asa, S.
J. Biol. Chem. 263, 13475-13478, 1988
A;Title: Glucagon gene expression in vertebrate brain.
A;Reference number: A30875; MUID:88330860; PMID:2901414
A;Accession: A30875
A;Molecule type: mRNA
A;Residues: 1-180 <DRU>
A;Cross-references: GB:J04040; NID:g183269; PIDN:AAA52567.1; PID:g183270
R;Orskov, C.; Bersani, M.; Johnsen, A.H.; Hojrup, P.; Holst, J.J.
J. Biol. Chem. 264, 12826-12829, 1989
A;Title: Complete sequences of glucagon-like peptide-1 from human and pig small intestin
A;Reference number: A92732; MUID:89327238; PMID:2753890
A;Accession: A32614
A;Molecule type: protein
A;Residues: 98-127 <ORS>
R;Thomsen, J.; Kristiansen, K.; Brunfeldt, K.; Sundby, F.
FEBS Lett. 21, 315-319, 1972
A;Title: The amino acid sequence of human glucagon.
A;Reference number: A91373
A;Accession: A01541
A;Molecule type: protein
A;Residues: 53-81 <THO>
R;Tsugita, A.; Takamoto, K.; Kamo, M.; Iwadate, H.
Eur. J. Biochem. 206, 691-696, 1992
A;Title: C-terminal sequencing of protein. A novel partial acid hydrolysis and analysis
A;Reference number: S23188; MUID:92298996; PMID:1606956
A;Accession: S23309
A;Molecule type: protein
A;Residues: 53-81 <TSU>
C;Comment: In pancreatic alpha-cells, proglucagon is processed to glicentin-related poly
stinal L cells, proglucagon is processed to truncated glucagon-like peptide 1, glucagon-
dulin.
C;Genetics:
A;Gene: GDB:GCG
A;Cross-references: GDB:119265; OMIM:138030
A;Map position: 2q36-2q37
A;Introns: 31/2; 85/2; 131/2; 179/2
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; int
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status experimental <PGC>
F;21-89/Product: glicentin #status experimental <GLN>
F;21-50/Product: glicentin-related polypeptide #status predicted <GRPP>
F;53-89/Product: oxyntomodulin #status experimental <OXN>
F;53-81/Product: glucagon #status experimental <GCN>
F;92-178/Product: major proglucagon fragment #status experimental <MPGF>
F;92-127/Product: glucagon-like peptide 1 #status experimental <GL1>
F;98-127/Product: truncated glucagon-like peptide 1 #status predicted <TGL>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 6.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 3
GCGP
glucagon precursor - guinea pig
N;Alternate names: oxyntomodulin
N;Contains: glicentin-related peptide; glucagon; glucagon-37 (oxyntomodulin); glucago
C;Species: Cavia porcellus (guinea pig)
C;Date: 30-Sep-1987 #sequence_revision 31-Dec-1992 #text_change 16-Jun-2000
C;Accession: A24856; A23849; A60323
R;Seino, S.; Welsh, M.; Bell, G.I.; Chan, S.J.; Steiner, D.F.
FEBS Lett. 203, 25-30, 1986
A;Title: Mutations in the guinea pig preproglucagon gene are restricted to a specific
A;Reference number: A24856; MUID:86248118; PMID:3755107
A;Accession: A24856
A;Molecule type: mRNA
A;Residues: 1-180 <SEI>
A;Cross-references: DDBJ:D00014; GB:N00014; NID:g220288; PIDN:BAA00010.1; PID:g220289
R;Huang, C.G.; Eng, J.; Pan, Y.C.E.; Hulmes, J.D.; Yalow, R.S.
Diabetes 35, 508-512, 1986
A;Title: Guinea pig glucagon differs from other mammalian glucagons.
A;Reference number: A23849; MUID:86165412; PMID:3956884
A;Accession: A23849
A;Molecule type: protein
A;Residues: 53-81 <HUA>
R;Conlon, J.M.; Hansen, H.F.; Schwartz, T.W.
Regul. Pept. 11, 309-320, 1985
A;Title: Primary structure of glucagon and a partial sequence of oxyntomodulin (gluca
A;Reference number: A60323; MUID:86017849; PMID:4048553
A;Accession: A60323
A;Molecule type: protein
A;Residues: 53-81 <CON>
A;Note: glucagon-37 was not completely sequenced
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pan
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-89/Product: glucagon-37 (oxyntomodulin) #status experimental <G37>
F;53-81/Product: glucagon #status experimental <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 6.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127
```

```
RESULT 4
GCRTDU
glucagon precursor - degu
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-like
C;Species: Octodon degus (degu)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: C36118
R;Nishi, M.; Steiner, D.F.
Mol. Endocrinol. 4, 1192-1198, 1990
A;Title: Cloning of complementary DNAs encoding islet amyloid polypeptide, insulin, and
A;Reference number: A36118; MUID:91155952; PMID:2293024
A;Accession: C36118
A;Molecule type: mRNA
A;Residues: 1-180 <NIS>
A;Cross-references: GB:M57688; NID:g202467; PIDN:AAA40588.1; PID:g202468
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancre
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status predicted <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 6.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 5
GCRT
glucagon precursor - rat
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-like
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 26-Feb-1999
C;Accession: A22655; A25190; A44198
R;Heinrich, G.; Gros, P.; Habener, J.F.
J. Biol. Chem. 259, 14082-14087, 1984
A;Title: Glucagon gene sequence: four of six exons encode separate functional domains of
A;Reference number: A22655; MUID:85054853; PMID:6094539
A;Accession: A22655
A;Molecule type: DNA
A;Residues: 1-180 <HEI>
A;Cross-references: EMBL:K02809
A;Note: the authors translated the codon TTG for residue 10 as Glu and ACC for residue 5
R;Mojsov, S.; Heinrich, G.; Wilson, I.B.; Ravazzola, M.; Orci, L.; Habener, J.F.
J. Biol. Chem. 261, 11880-11889, 1986
A;Title: Preproglucagon gene expression in pancreas and intestine diversifies at the lev
A;Reference number: A25190; MUID:86304324; PMID:3528148
A;Accession: A25190
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-180 <MOJ>
R;Heinrich, G.; Gros, P.; Lund, P.K.; Bentley, R.C.; Habener, J.F.
Endocrinology 115, 2176-2181, 1984
A;Title: Pre-proglucagon messenger ribonucleic acid: nucleotide and encoded amino acid s
A;Reference number: A44198; MUID:85051023; PMID:6548696
A;Accession: A44198
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-180 <HE2>
A;Cross-references: GB:K02809; GB:K02810; GB:K02811; GB:K02812
C;Genetics:
A;Introns: 31/2; 85/2; 131/2; 179/2
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancre
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status predicted <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-180/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 6.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 6
GCHY
glucagon precursor - golden hamster
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-li
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 20-Mar-1998
C;Accession: A01539
R;Bell, G.I.; Santerre, R.F.; Mullenbach, G.T.
Nature 302, 716-718, 1983
A;Title: Hamster preproglucagon contains the sequence of glucagon and two related pep
A;Reference number: A01539; MUID:83167563; PMID:6835407
A;Accession: A01539
A;Molecule type: mRNA
A;Residues: 1-180 <BEL>
A;Cross-references: EMBL:J00059
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pan
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status predicted <GCN>
F;98-127/Product: glucagon-like peptide 1 #status predicted <GL1>
F;146-180/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 6.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 7
GCBO
glucagon precursor - bovine
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1; glucagon-li
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 20-Mar-1998
C;Accession: A93970; A92081; A01538
R;Lopez, L.C.; Frazier, M.L.; Su, C.J.; Kumar, A.; Saunders, G.F.
Proc. Natl. Acad. Sci. U.S.A. 80, 5485-5489, 1983
A;Title: Mammalian pancreatic preproglucagon contains three glucagon-related peptides
A;Reference number: A93970; MUID:83299996; PMID:6577439
A;Accession: A93970
A;Molecule type: mRNA
A;Residues: 1-180 <LOP>
A;Cross-references: EMBL:K00107
R;Bromer, W.W.; Boucher, M.E.; Koffenberger Jr., J.E.
J. Biol. Chem. 246, 2822-2827, 1971
A;Title: Amino acid sequence of bovine glucagon.
A;Reference number: A92081; MUID:71166445; PMID:5102927
A;Accession: A92081
A;Molecule type: protein
A;Residues: 53-81 <BRO>
C;Superfamily: glucagon
```

```
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pancre
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-180/Product: proglucagon #status predicted <PGC>
F;21-50/Region: glicentin-related peptide #status predicted
F;53-81/Product: glucagon #status experimental <GCN>
F;98-127/Product: glucagon-like peptide 1 #status experimental <GL1>
F;146-178/Product: glucagon-like peptide 2 #status predicted <GL2>
F;127/Modified site: amidated carboxyl end (Arg) (amide in mature form from following gl

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 6.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db     98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 8
A57294
glucagon precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
C;Accession: A57294; S49903
R;Rothenberg, M.E.; Eilertson, C.D.; Klein, K.; Zhou, Y.; Lindberg, I.; McDonald, J.K.;
J. Biol. Chem. 270, 10136-10146, 1995
A;Title: Processing of mouse proglucagon by recombinant prohormone convertase 1 and immu
A;Reference number: A57294; MUID:95247722; PMID:7730317
A;Accession: A57294
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-180 <ROT>
A;Cross-references: EMBL:Z46845; NID:g599880; PIDN:CAA86902.1; PID:g599881
C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas

  Query Match             94.3%;  Score 133;  DB 2;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 6.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db     98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 9
GCCH
glucagon precursor - chicken
N;Contains: glucagon; glucagon-like peptide 1
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1991 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: S09992; A92189; A60836; A01542
R;Hasegawa, S.; Terazono, K.; Nata, K.; Takada, T.; Yamamoto, H.; Okamoto, H.
FEBS Lett. 264, 117-120, 1990
A;Title: Nucleotide sequence determination of chicken glucagon precursor cDNA. Chicken p
A;Reference number: S09992; MUID:90249492; PMID:2338135
A;Accession: S09992
A;Molecule type: mRNA
A;Residues: 1-151 <HAS>
A;Cross-references: EMBL:Y07539; NID:g63749; PIDN:CAA68827.1; PID:g63750
R;Pollock, H.G.; Kimmel, J.R.
J. Biol. Chem. 250, 9377-9380, 1975
A;Title: Chicken glucagon. Isolation and amino acid sequence studies.
A;Reference number: A92189; MUID:76069271; PMID:1194290
A;Accession: A92189
A;Molecule type: protein
A;Residues: 55-83 <POL>
R;Huang, J.; Eng, J.; Yalow, R.S.
Horm. Metab. Res. 19, 542-544, 1987
A;Title: Chicken glucagon: sequence and potency in receptor assay.
A;Reference number: A60836; MUID:88113418; PMID:2828209
A;Accession: A60836
A;Molecule type: protein
A;Residues: 55-83 <HUA>
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; carbohydrate metabolism; duplication; hormone; pan
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-151/Product: proglucagon #status predicted <PGC>
F;55-83/Product: glucagon #status experimental <GCN>
F;118-147/Product: glucagon-like peptide 1 #status predicted <GL1>
F;147/Modified site: amidated carboxyl end (Arg) (amide in mature form from following

  Query Match             85.8%;  Score 121;  DB 1;  Length 151;
  Best Local Similarity   73.3%;  Pred. No. 4.5e-12;
  Matches   22;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | |||:|||::||| ||||  |||||| ||
Db    118 HAEGTYTSDITSYLEGQAAKEFIAWLVNGR 147


RESULT 10
I51301
proglucagon - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51301
R;Irwin, D.M.; Wong, J.
Mol. Endocrinol. 9, 267-277, 1995
A;Title: Trout and chicken proglucagon: alternative splicing generates mRNA transcrip
A;Reference number: A55895; MUID:95295739; PMID:7776976
A;Accession: I51301
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-206 <IRW>
A;Cross-references: GB:S78477; NID:g999386; PIDN:AAB34506.1; PID:g999387
C;Superfamily: glucagon
C;Keywords: duplication

  Query Match             85.8%;  Score 121;  DB 2;  Length 206;
  Best Local Similarity   73.3%;  Pred. No. 6.2e-12;
  Matches   22;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | |||:|||::||| ||||  |||||| ||
Db    118 HAEGTYTSDITSYLEGQAAKEFIAWLVNGR 147


RESULT 11
B61125
glucagon-like peptide - American eel
C;Species: Anguilla rostrata (American eel)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 21-Nov-1997
C;Accession: B61125
R;Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
Gen. Comp. Endocrinol. 82, 23-32, 1991
A;Title: The primary structure of glucagon-like peptide but not insulin has been cons
A;Reference number: A61125; MUID:91340068; PMID:1874385
A;Accession: B61125
A;Molecule type: protein
A;Residues: 1-30 <CON>
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; duplication
F;1-30/Product: glucagon-like peptide #status experimental <GLP>
F;30/Modified site: amidated carboxyl end (Arg) #status predicted

  Query Match             75.9%;  Score 107;  DB 2;  Length 30;
  Best Local Similarity   66.7%;  Pred. No. 1.5e-10;
  Matches   20;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | |||:||||||||  |||  |::||  ||
Db      1 HAEGTYTSDVSSYLQDQAAKEFVSWLKTGR 30


RESULT 12
```

```
C61125
glucagon-like peptide - European eel
C;Species: Anguilla anguilla (European eel)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 21-Nov-1997
C;Accession: C61125
R;Conlon, J.M.; Andrews, P.C.; Thim, L.; Moon, T.W.
Gen. Comp. Endocrinol. 82, 23-32, 1991
A;Title: The primary structure of glucagon-like peptide but not insulin has been conserv
A;Reference number: A61125; MUID:91340068; PMID:1874385
A;Accession: C61125
A;Molecule type: protein
A;Residues: 1-30 <CON>
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; duplication
F;1-30/Product: glucagon-like peptide #status experimental <GLP>
F;30/Modified site: amidated carboxyl end (Arg) #status experimental

  Query Match             75.9%;  Score 107;  DB 2;  Length 30;
  Best Local Similarity   66.7%;  Pred. No. 1.5e-10;
  Matches   20;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | |||:||||||||  |||  |::||  ||
Db       1 HAEGTYTSDVSSYLQDQAAKEFVSWLKTGR 30


RESULT 13
GCFGB
glucagon precursor - bullfrog (fragments)
N;Alternate names: oxyntomodulin
N;Contains: glucagon; glucagon-36 (oxyntomodulin); glucagon-like peptide 1; glucagon-lik
C;Species: Rana catesbeiana (bullfrog)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C;Accession: B28091; C28091; D28091
R;Pollock, H.G.; Hamilton, J.W.; Rouse, J.B.; Ebner, K.E.; Rawitch, A.B.
J. Biol. Chem. 263, 9746-9751, 1988
A;Title: Isolation of peptide hormones from the pancreas of the bullfrog (Rana catesbeia
A;Reference number: A92730; MUID:88257102; PMID:3260236
A;Accession: B28091
A;Molecule type: protein
A;Residues: 1-36 <PO2>
A;Accession: C28091
A;Molecule type: protein
A;Residues: 37-68 <POL>
A;Accession: D28091
A;Molecule type: protein
A;Residues: 69-101 <PO3>
C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas
F;1-36/Product: glucagon-36 (oxyntomodulin) #status experimental <G36>
F;1-29/Product: glucagon #status predicted <GCN>
F;37-67/Product: glucagon-like peptide 1 #status experimental <GL1>
F;69-101/Product: glucagon-like peptide 2 #status experimental <GL2>

  Query Match             75.9%;  Score 107;  DB 1;  Length 101;
  Best Local Similarity   63.3%;  Pred. No. 5.4e-10;
  Matches   19;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | :||||||:||||  :||  |: ||:|||
Db      37 HADGTFTSDMSSYLEEKAAKEFVDWLIKGR 66


RESULT 14
GCAF2
glucagon 2 precursor - American goosefish
N;Contains: glucagon; glucagon-like peptide 1
C;Species: Lophius americanus (American goosefish)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 21-Jul-2000
C;Accession: A05150
R;Lund, P.K.; Goodman, R.H.; Montminy, M.R.; Dee, P.C.; Habener, J.F.
J. Biol. Chem. 258, 3280-3284, 1983
A;Title: Anglerfish islet pre-proglucagon II. Nucleotide and corresponding amino acid se
A;Reference number: A05150; MUID:83135785; PMID:6338015
A;Accession: A05150
A;Molecule type: mRNA
A;Residues: 1-122 <LUN>
A;Cross-references: GB:J00933; NID:g64021; PIDN:CAA23905.1; PID:g64022
C;Superfamily: glucagon
C;Keywords: carbohydrate metabolism; duplication; hormone; pancreas
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-122/Product: proglucagon 2 #status predicted <PGC2>
F;52-80/Product: glucagon #status predicted <GCN>
F;89-119/Product: glucagon-like peptide 1 #status predicted <GL1>

  Query Match             73.8%;  Score 104;  DB 1;  Length 122;
  Best Local Similarity   63.3%;  Pred. No. 2e-09;
  Matches   19;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | :||:||||||||  |||  |::||  ||
Db      89 HADGTYTSDVSSYLQDQAAKDFVSWLKAGR 118


RESULT 15
I51093
glucagon - chinook salmon (fragment)
C;Species: Oncorhynchus tschawytscha (chinook salmon)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51093
R;Irwin, D.M.; Wong, J.
Mol. Endocrinol. 9, 267-277, 1995
A;Title: Trout and chicken proglucagon: alternative splicing generates mRNA transcrip
A;Reference number: A55895; MUID:95295739; PMID:7776976
A;Accession: I51093
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-66 <IRW>
A;Cross-references: EMBL:U19920; NID:g736366; PIDN:AAC59670.1; PID:g736367
C;Superfamily: glucagon
C;Keywords: duplication

  Query Match             72.3%;  Score 102;  DB 2;  Length 66;
  Best Local Similarity   60.0%;  Pred. No. 2.2e-09;
  Matches   18;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | :||:|||||:||  |||  |::||  ||
Db      33 HADGTYTSDVSTYLQDQAAKDFVSWLKSGR 62


Search completed: February 13, 2003, 11:03:32
Job time : 17 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 13, 2003, 10:58:31 ; Search time 11 Seconds
(without alignments)
116.888 Million cell updates/sec

Title: US-09-091-605-1
Perfect score: 141
Sequence: 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	133	94.3	158	1	GLUC_PIG	P01274 sus scrofa
2	133	94.3	180	1	GLUC_BOVIN	P01272 bos taurus
3	133	94.3	180	1	GLUC_CAVPO	P05110 cavia porce
4	133	94.3	180	1	GLUC_HUMAN	P01275 homo sapien
5	133	94.3	180	1	GLUC_MESAU	P01273 mesocricetu
6	133	94.3	180	1	GLUC_MOUSE	P55095 mus musculu
7	133	94.3	180	1	GLUC_OCTDE	P22890 octodon deg
8	133	94.3	180	1	GLUC_RAT	P06883 rattus norv
9	121	85.8	151	1	GLUC_CHICK	P01277 gallus gall
10	107	75.9	30	1	GLUM_ANGAN	P41521 anguilla an
11	107	75.9	103	1	GLUC_RANCA	P15438 rana catesb
12	104	73.8	122	1	GLU2_LOPAM	P04092 lophius ame
13	102	72.3	121	1	GLUC_CARAU	P79695 carassius a
14	101	71.6	78	1	GLUC_LEPSP	P09566 lepisosteus
15	100	70.9	71	1	GLUC_ICTPU	P04093 ictalurus p
16	98	69.5	71	1	GLUC_PIAME	P81880 piaractus m
17	97	68.8	68	1	GLUC_ONCKI	P07449 oncorhynchu
18	94	66.7	33	1	GLU2_ORENI	P81027 oreochromis
19	89	63.1	96	1	GLUC_MYOSC	P09686 myoxocephal
20	84	59.6	29	1	GLUC_TORMA	P09567 torpedo mar
21	82	58.2	29	1	GLUC_SCYCA	P09687 scyliorhinu
22	81	57.4	29	1	GLUC_DIDMA	P18108 didelphis m
23	81	57.4	29	1	GLUC_RABIT	P25449 oryctolagus
24	81	57.4	69	1	GLUC_CANFA	P29794 canis famil
25	81	57.4	124	1	GLU1_LOPAM	P01278 lophius ame
26	80	56.7	29	1	GLUC_CALMI	P13189 callorhynch
27	79	56.0	29	1	GLUC_ANAPL	P01276 anas platyr
28	79	56.0	87	1	EXE4_HELSU	P26349 heloderma s
29	78	55.3	36	1	GLU1_ORENI	P81026 oreochromis
30	77	54.6	29	1	GLUC_LAMFL	Q9prq9 lampetra fl
31	77	54.6	39	1	EXE3_HELHO	P20394 heloderma h
32	75	53.2	29	1	GLUC_CHIBR	P31297 chinchilla
33	74	52.5	29	1	GLUC_PLAFE	P23062 platichthys
34	74	52.5	75	1	GLUC_AMICA	P33528 amia calva
35	67	47.5	36	1	GLUC_HYDCO	P09682 hydrolagus
36	49	34.8	42	1	GIP_BOVIN	P09680 bos taurus
37	49	34.8	42	1	GIP_PIG	P01281 sus scrofa
38	49	34.8	144	1	GIP_MOUSE	P48756 mus musculu
39	49	34.8	144	1	GIP_RAT	Q06145 rattus norv
40	48.5	34.4	326	1	UPK_SYNY3	Q55684 synechocyst
41	48	34.0	72	1	VIP_BOVIN	P81401 bos taurus
42	48	34.0	72	1	VIP_CAVPO	P04566 cavia porce
43	48	34.0	72	1	VIP_PIG	P01284 sus scrofa
44	48	34.0	72	1	VIP_RABIT	P32649 oryctolagus
45	48	34.0	153	1	GIP_HUMAN	P09681 homo sapien

ALIGNMENTS

RESULT 1
GLUC_PIG
ID GLUC_PIG STANDARD; PRT; 158 AA.
AC P01274;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucagon precursor [Contains: Glicentin; Glicentin-related polypeptide
DE (GRPP); Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like
DE peptide 2 (GLP2)] (Fragment).
GN GCG.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE OF 1-69.
RX MEDLINE=81248172; PubMed=6894800;
RA Thim L., Moody A.J.;
RT "The primary structure of porcine glicentin (proglucagon).";
RL Regul. Pept. 2:139-150(1981).
RN [2]
RP SEQUENCE OF 1-69.
RX MEDLINE=82221776; PubMed=7045833;
RA Thim L., Moody A.J.;
RT "The amino acid sequence of porcine glicentin.";
RL Peptides 2 Suppl. 2:37-39(1981).
RN [3]
RP SEQUENCE OF 33-61.
RA Bromer W.W., Sinn L.G., Behrens O.K.;
RT "The amino acid sequence of glucagon. V. Location of amide groups,
RT acid degradation studies and summary of sequential evidence.";
RL J. Am. Chem. Soc. 79:2807-2810(1957).
RN [4]
RP SEQUENCE OF 78-107.
RX MEDLINE=89327238; PubMed=2753890;
RA Orskov C., Bersani M., Johnsen A.H., Hoejrup P., Holst J.J.;
RT "Complete sequences of glucagon-like peptide-1 from human and pig
RT small intestine.";
RL J. Biol. Chem. 264:12826-12829(1989).
RN [5]
RP SEQUENCE OF 111-158.
RX MEDLINE=88243712; PubMed=3379036;
RA Buhl T., Thim L., Kofod H., Orskov C., Harling H., Holst J.J.;
RT "Naturally occurring products of proglucagon 111-160 in the porcine
RT and human small intestine.";
RL J. Biol. Chem. 263:8621-8624(1988).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=76051297; PubMed=171582;
RA Sasaki K., Dockerill S., Adamiak D.A., Tickle I.J., Blundell T.L.;
RT "X-ray analysis of glucagon and its relationship to receptor
RT binding.";
RL Nature 257:751-757(1975).
CC -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC RAISES THE BLOOD SUGAR LEVEL.

```
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- MISCELLANEOUS: X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH
CC       HUMAN SEQUENCE.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   PIR; A01540; GCPG.
DR   PDB; 1GCN; 30-SEP-83.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 3.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues;
KW   3D-structure.
FT   NON_TER       1      1
FT   PEPTIDE       1     69       GLICENTIN.
FT   PEPTIDE       1     30       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      33     61       GLUCAGON.
FT   PEPTIDE      78    107       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     126    158       GLUCAGON-LIKE PEPTIDE 2.
FT   HELIX        39     42
FT   TURN         43     45
FT   HELIX        46     55
FT   TURN         56     57
SQ   SEQUENCE   158 AA;  18212 MW;  28C6FCF257F333B2 CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 158;
  Best Local Similarity   86.7%;  Pred. No. 2.6e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      78 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 107


RESULT 2
GLUC_BOVIN
ID   GLUC_BOVIN     STANDARD;      PRT;   180 AA.
AC   P01272;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83299996; PubMed=6577439;
RA   Lopez L.C., Frazier M.L., Su C.-J., Kumar A., Saunders G.F.;
RT   "Mammalian pancreatic preproglucagon contains three glucagon-related
RT   peptides.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:5485-5489(1983).
RN   [2]
RP   SEQUENCE OF 53-81.
RX   MEDLINE=71166445; PubMed=5102927;
RA   Bromer W.W., Boucher M.E., Koffenberger J.E. Jr.;
RT   "Amino acid sequence of bovine glucagon.";
RL   J. Biol. Chem. 246:2822-2827(1971).
RN   [3]
RP   STRUCTURE BY NMR OF 53-81.
RX   MEDLINE=71166445; PubMed=6631957;
RA   Braun W., Wider G., Lee K.H., Wuthrich K.;
RT   "Conformation of glucagon in a lipid-water interphase by 1H nuclear
RT   magnetic resonance.";
RL   J. Mol. Biol. 169:921-948(1983).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K00107; AAA30538.1; -.
DR   PIR; A01538; GCBO.
DR   PDB; 1KX6; 13-FEB-02.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   3D-structure.
FT   SIGNAL        1     20
FT   PEPTIDE      21     50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE      92    128       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     146    178       GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20944 MW;  8D9B4FF05B9F15FF CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 3
GLUC_CAVPO
ID   GLUC_CAVPO     STANDARD;      PRT;   180 AA.
AC   P05110;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-37 (Oxyntomodulin); Glucagon-like peptide 1 (GLP1);
DE   Glucagon-like peptide 2 (GLP2)].
GN   GCG.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86248118; PubMed=3755107;
RA   Seino S., Welsh M., Bell G.I., Chan S.J., Steiner D.F.;
RT   "Mutations in the guinea pig preproglucagon gene are restricted to a
RT   specific portion of the prohormone sequence.";
RL   FEBS Lett. 203:25-30(1986).
RN   [2]
RP   SEQUENCE OF 53-81.
RX   MEDLINE=86165412; PubMed=3956884;
RA   Huang C.G., Eng J., Pan Y.-C.E., Hulmes J.D., Yalow R.S.;
RT   "Guinea pig glucagon differs from other mammalian glucagons.";
RL   Diabetes 35:508-512(1986).
RN   [3]
RP   PARTIAL SEQUENCE OF 53-89.
```

```
RX   MEDLINE=86017849; PubMed=4048553;
RA   Conlon J.M., Hansen H.F., Schwartz T.W.;
RT   "Primary structure of glucagon and a partial sequence of
RT   oxyntomodulin (glucagon-37) from the guinea pig.";
RL   Regul. Pept. 11:309-320(1985).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00014; BAA00010.1; -.
DR   PIR; A24856; GCGP.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL        1     20
FT   PEPTIDE      21     50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE      53     89       GLUCAGON-37.
FT   PEPTIDE      92    128       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     146    178       GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20972 MW;  702FB181161D2776 CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 4
GLUC_HUMAN
ID   GLUC_HUMAN     STANDARD;      PRT;   180 AA.
AC   P01275;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88330860; PubMed=2901414;
RA   Drucker D.J., Asa S.;
RT   "Glucagon gene expression in vertebrate brain.";
RL   J. Biol. Chem. 263:13475-13478(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86259053; PubMed=3725587;
RA   White J.W., Saunders G.F.;
RT   "Structure of the human glucagon gene.";
RL   Nucleic Acids Res. 14:4719-4730(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=83271477; PubMed=6877358;
RA   Bell G.I., Sanchez-Pescador R., Laybourn P.J., Najarian R.C.;
RT   "Exon duplication and divergence in the human preproglucagon gene.";
RL   Nature 304:368-371(1983).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 53-81.
RA   Thomsen J., Kristiansen K., Brunfeldt K., Sundby F.;
RT   "The amino acid sequence of human glucagon.";
RL   FEBS Lett. 21:315-319(1972).
RN   [6]
RP   SEQUENCE OF 98-127.
RX   MEDLINE=89327238; PubMed=2753890;
RA   Orskov C., Bersani M., Johnsen A.H., Hoejrup P., Holst J.J.;
RT   "Complete sequences of glucagon-like peptide-1 from human and pig
RT   small intestine.";
RL   J. Biol. Chem. 264:12826-12829(1989).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 53-81.
RX   MEDLINE=98334683; PubMed=9667960;
RA   Sturm N.S., Lin Y., Burley S.K., Krstenansky J.L., Ahn J.M.,
RA   Azizeh B.Y., Trivedi D., Hruby V.J.;
RT   "Structure-function studies on positions 17, 18, and 21 replacement
RT   analogues of glucagon: the importance of charged residues and salt
RT   bridges in glucagon biological activity.";
RL   J. Med. Chem. 41:2693-2700(1998).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- PHARMACEUTICAL: Available under the names Glucagon (Eli Lilly) and
CC       GlucaGen or Glucagon Novo Nordisk (Novo Nordisk). Used to treat
CC       severe hypoglycemia in insulin-dependent diabetics.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   -!- DATABASE: NAME=Glucagon at Eli Lilly;
CC       NOTE=Clinical information on Eli Lilly glucagon products;
CC       WWW="http://www.lillyDiabetes.com/Products/PatientInfo.cfm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04040; AAA52567.1; -.
DR   EMBL; X03991; CAA27627.1; -.
DR   EMBL; V01515; CAA24759.1; -.
DR   EMBL; BC005278; AAH05278.1; -.
DR   PIR; A24377; GCHU.
DR   PIR; S23309; S23309.
DR   PDB; 1BH0; 18-NOV-98.
DR   Genew; HGNC:4191; GCG.
DR   MIM; 138030; -.
DR   MIM; 231530; -.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
```

```
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Pharmaceutical; 3D-structure.
FT   SIGNAL       1     20
FT   PEPTIDE      21    50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53    81       GLUCAGON.
FT   PEPTIDE      98    127      GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE      146   178      GLUCAGON-LIKE PEPTIDE 2.
FT   CONFLICT     82    82       K -> N (IN REF. 3).
SQ   SEQUENCE   180 AA;  20909 MW;  7A99EEC629B2862C CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 5
GLUC_MESAU
ID   GLUC_MESAU     STANDARD;      PRT;   180 AA.
AC   P01273;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83167563; PubMed=6835407;
RA   Bell G.I., Santerre R.F., Mullenbach G.T.;
RT   "Hamster preproglucagon contains the sequence of glucagon and two
RT   related peptides.";
RL   Nature 302:716-718(1983).
RN   [2]
RP   REVISIONS TO 12-15.
RA   Bell G.I.;
RL   Submitted (XXX-1985) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J00059; AAA37074.1; -.
DR   PIR; A01539; GCHY.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL       1     20
```

```
FT   PEPTIDE      21    50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53    81       GLUCAGON.
FT   PEPTIDE      92    128      GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE      146   178      GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20954 MW;  02791B49D7AADD4B CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 6
GLUC_MOUSE
ID   GLUC_MOUSE     STANDARD;      PRT;   180 AA.
AC   P55095;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreatic islets;
RX   MEDLINE=95247722; PubMed=7730317;
RA   Rothenberg M.E., Eilertson C.D., Klein K., Zhou Y., Linberg I.,
RA   McDonald J.K., Mackin R.B., Noe B.D.;
RT   "Processing of mouse proglucagon by recombinant prohormone convertase
RT   1 and immunopurified prohormone convertase 2 in vitro.";
RL   J. Biol. Chem. 270:10136-10146(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Shamsadin R., Knepel W.;
RT   "Mouse glucagon full length cDNA.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z46845; CAA86902.1; -.
DR   EMBL; AF276754; AAK96898.1; -.
DR   HSSP; P01274; 1GCN.
DR   MGD; MGI:95674; Gcg.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL       1     20       BY SIMILARITY.
FT   PEPTIDE      21    50       GLICENTIN-RELATED POLYPEPTIDE.
```

```
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE      92    128       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     146    178       GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20906 MW;  595AA6DD9A589950 CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 7
GLUC_OCTDE
ID   GLUC_OCTDE     STANDARD;      PRT;   180 AA.
AC   P22890;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Octodon degus (Degu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Octodontidae; Octodon.
OX   NCBI_TaxID=10160;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91155952; PubMed=2293024;
RA   Nishi M., Steiner D.F.;
RT   "Cloning of complementary DNAs encoding islet amyloid polypeptide,
RT   insulin, and glucagon precursors from a New World rodent, the degu,
RT   Octodon degus.";
RL   Mol. Endocrinol. 4:1192-1198(1990).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57688; AAA40588.1; -.
DR   PIR; C36118; GCRTDU.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Amidation.
FT   SIGNAL        1     20
FT   PEPTIDE      21     50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE      92    127       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     146    178       GLUCAGON-LIKE PEPTIDE 2.
FT   MOD_RES     127    127       AMIDATION (G-128 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   180 AA;  21165 MW;  6E8836160A9A3051 CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 8
GLUC_RAT
ID   GLUC_RAT       STANDARD;      PRT;   180 AA.
AC   P06883;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide 1 (GLP1); Glucagon-like peptide 2
DE   (GLP2)].
GN   GCG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85054853; PubMed=6094539;
RA   Heinrich G., Gros P., Habener J.F.;
RT   "Glucagon gene sequence. Four of six exons encode separate functional
RT   domains of rat pre-proglucagon.";
RL   J. Biol. Chem. 259:14082-14087(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85051023; PubMed=6548696;
RA   Heinrich G., Gros P., Lund P.K., Bentley R.C., Habener J.F.;
RT   "Pre-proglucagon messenger ribonucleic acid: nucleotide and encoded
RT   amino acid sequences of the rat pancreatic complementary
RT   deoxyribonucleic acid.";
RL   Endocrinology 115:2176-2181(1984).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86304324; PubMed=3528148;
RA   Mojsov S., Heinrich G., Wilson I.B., Ravazzola M., Orci L.,
RA   Habener J.F.;
RT   "Preproglucagon gene expression in pancreas and intestine diversifies
RT   at the level of post-translational processing.";
RL   J. Biol. Chem. 261:11880-11889(1986).
CC   -!- FUNCTION: GLUCAGON PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND
CC       RAISES THE BLOOD SUGAR LEVEL.
CC   -!- FUNCTION: GLP2 STIMULATES INTESTINAL GROWTH AND UPREGULATES VILLUS
CC       HEIGHT IN THE SMALL INTESTINE, CONCOMITANT WITH INCREASED CRYPT
CC       CELL PROLIFERATION AND DECREASED ENTEROCYTE APOPTOSIS.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02813; AAA41235.1; -.
DR   EMBL; K02809; AAA41235.1; JOINED.
DR   EMBL; K02810; AAA41235.1; JOINED.
DR   EMBL; K02811; AAA41235.1; JOINED.
DR   EMBL; K02812; AAA41235.1; JOINED.
DR   PIR; A22655; GCRT.
DR   PIR; A44198; A44198.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
```

```
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 4.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL        1     20
FT   PEPTIDE      21     50       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE      92    128       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     146    178       GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   180 AA;  20846 MW;  76931409D03C7978 CRC64;

  Query Match             94.3%;  Score 133;  DB 1;  Length 180;
  Best Local Similarity   86.7%;  Pred. No. 3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127


RESULT 9
GLUC_CHICK
ID   GLUC_CHICK     STANDARD;      PRT;   151 AA.
AC   P01277;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Glucagon precursor.
OS   Gallus gallus (Chicken), and
OS   Meleagris gallopavo (Common turkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031, 9103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Chicken; TISSUE=Pancreas;
RX   MEDLINE=90249492; PubMed=2338135;
RA   Hasegawa S., Terazono K., Nata K., Takada T., Yamamoto H.,
RA   Okamoto H.;
RT   "Nucleotide sequence determination of chicken glucagon precursor
RT   cDNA. Chicken preproglucagon does not contain glucagon-like peptide
RT   II.";
RL   FEBS Lett. 264:117-120(1990).
RN   [2]
RP   SEQUENCE OF 55-83.
RC   SPECIES=Chicken;
RX   MEDLINE=76069271; PubMed=1194290;
RA   Pollock H.G., Kimmel J.R.;
RT   "Chicken glucagon. Isolation and amino acid sequence studies.";
RL   J. Biol. Chem. 250:9377-9380(1975).
RN   [3]
RP   COMPOSITION, AND SEQUENCE OF 55-83.
RC   SPECIES=M.gallopavo;
RX   MEDLINE=73074118; PubMed=4645932;
RA   Markussen J., Frandsen E.K., Heding L.G., Sundby F.;
RT   "Turkey glucagon: crystallization, amino acid composition and
RT   immunology.";
RL   Horm. Metab. Res. 4:360-363(1972).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- MISCELLANEOUS: THE COMPOSITION OF TURKEY GLUCAGON APPEARS TO BE
CC       IDENTICAL WITH CHICKEN.
CC   -!- MISCELLANEOUS: CHICKEN PREPROGLUCAGON DOES NOT CONTAIN
CC       GLUCAGON-LIKE PEPTIDE II.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y07539; CAA68827.1; -.
DR   PIR; S09992; GCCH.
DR   PIR; A91740; A91740.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 3.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Amidation.
FT   SIGNAL        1     22
FT   CHAIN        23    151       PROGLUCAGON.
FT   PEPTIDE      55     83       GLUCAGON.
FT   PROPEP       86    118
FT   PEPTIDE     118    147       GLUCAGON-LIKE PEPTIDE.
FT   MOD_RES      147    147       AMIDATION (G-148 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   151 AA;  17520 MW;  B6C0D87536C0AEB5 CRC64;

  Query Match             85.8%;  Score 121;  DB 1;  Length 151;
  Best Local Similarity   73.3%;  Pred. No. 2.1e-12;
  Matches   22;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | |||:|||::||| ||||  |||||| ||
Db     118 HAEGTYTSDITSYLEGQAAKEFIAWLVNGR 147


RESULT 10
GLUM_ANGAN
ID   GLUM_ANGAN     STANDARD;      PRT;    30 AA.
AC   P41521;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Glucagon-like peptide (GLP).
OS   Anguilla anguilla (European freshwater eel), and
OS   Anguilla rostrata (American eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC   Anguilla.
OX   NCBI_TaxID=7936, 7938;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pancreas;
RX   MEDLINE=91340068; PubMed=1874385;
RA   Conlon J.M., Andrews P.C., Thim L., Moon T.W.;
RT   "The primary structure of glucagon-like peptide but not insulin has
RT   been conserved between the American eel, Anguilla rostrata and the
RT   European eel, Anguilla anguilla.";
RL   Gen. Comp. Endocrinol. 82:23-32(1991).
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   PIR; B61125; B61125.
DR   PIR; C61125; C61125.
DR   HSSP; P01275; 1BH0.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 1.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 1.
DR   PROSITE; PS00260; GLUCAGON; 1.
KW   Glucagon family; Amidation.
FT   MOD_RES      30     30       AMIDATION.
SQ   SEQUENCE   30 AA;  3376 MW;  592DA5EABD6E49D0 CRC64;

  Query Match             75.9%;  Score 107;  DB 1;  Length 30;
  Best Local Similarity   66.7%;  Pred. No. 7.7e-11;
  Matches   20;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
```

```
          | |||:|||||||  |||  |::||  ||
Db      1 HAEGTYTSDVSSYLQDQAAKEFVSWLKTGR 30


RESULT 11
GLUC_RANCA
ID   GLUC_RANCA     STANDARD;      PRT;   103 AA.
AC   P15438; P15439; P15440;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   Glucagon precursor (Fragments).
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pancreas;
RX   MEDLINE=88257102; PubMed=3260236;
RA   Pollock H.G., Hamilton J.W., Rouse J.B., Ebner K.E., Rawitch A.B.;
RT   "Isolation of peptide hormones from the pancreas of the bullfrog
RT   (Rana catesbeiana). Amino acid sequences of pancreatic polypeptide,
RT   oxyntomodulin, and two glucagon-like peptides.";
RL   J. Biol. Chem. 263:9746-9751(1988).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- MISCELLANEOUS: X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH
CC       OTHER SPECIES SEQUENCES.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   PIR; B28091; GCFGB.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 3.
KW   Glucagon family; Hormone.
FT   PEPTIDE       1     29       GLUCAGON.
FT   PEPTIDE       1     36       GLUCAGON-36 (OXYNTOMODULIN).
FT   PEPTIDE      39     70       GLUCAGON-LIKE PEPTIDE 1.
FT   NON_CONS     70     71
FT   PEPTIDE      71    103       GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   103 AA;  11719 MW;  316287B7BAE1C8F7 CRC64;

  Query Match             75.9%;  Score 107;  DB 1;  Length 103;
  Best Local Similarity   63.3%;  Pred. No. 2.6e-10;
  Matches   19;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | :|||||:||||  :||  |: ||:|||
Db      39 HADGTFTSDMSSYLEEKAAKEFVDWLIKGR 68


RESULT 12
GLU2_LOPAM
ID   GLU2_LOPAM     STANDARD;      PRT;   122 AA.
AC   P04092;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucagon II precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon II; Glucagon-like peptide II].
OS   Lophius americanus (American goosefish) (Anglerfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.
OX   NCBI_TaxID=8073;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83135785; PubMed=6338015;
RA   Lund P.K., Goodman R.H., Montminy M.R., Dee P.C., Habener J.F.;
RT   "Anglerfish islet pre-proglucagon II. Nucleotide and corresponding
RT   amino acid sequence of the cDNA.";
RL   J. Biol. Chem. 258:3280-3284(1983).
RN   [2]
RP   PROCESSING.
RX   MEDLINE=86286913; PubMed=3526301;
RA   Noe B.D., Andrews P.C.;
RT   "Specific glucagon-related peptides isolated from anglerfish islets
RT   are metabolic cleavage products of (pre)proglucagon-II.";
RL   Peptides 7:331-339(1986).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00632; CAA23905.1; -.
DR   PIR; A05150; GCAF2.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL        1     21
FT   PEPTIDE      22     49       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      52     80       GLUCAGON II.
FT   PROPEP       83     86
FT   PEPTIDE      89    119       GLUCAGON-LIKE PEPTIDE II.
SQ   SEQUENCE   122 AA;  14171 MW;  5140AC47EF915519 CRC64;

  Query Match             73.8%;  Score 104;  DB 1;  Length 122;
  Best Local Similarity   63.3%;  Pred. No. 9.3e-10;
  Matches   19;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | :||:|||||||||  |||  |::||  ||
Db      89 HADGTYTSDVSSYLQDQAAKDFVSWLKAGR 118


RESULT 13
GLUC_CARAU
ID   GLUC_CARAU     STANDARD;      PRT;   121 AA.
AC   P79695;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucagon precursor [Contains: Glicentin-related polypeptide (GRPP);
DE   Glucagon; Glucagon-like peptide].
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yuen T.T.H., Mok P.Y., Chow B.K.C.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U65528; AAB39563.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal.
FT   SIGNAL       1     21       POTENTIAL.
FT   PEPTIDE      22    47       GLICENTIN-RELATED POLYPEPTIDE.
FT   PEPTIDE      50    78       GLUCAGON.
FT   PROPEP       80    85
FT   PEPTIDE      88    121      GLUCAGON-LIKE PEPTIDE.
SQ   SEQUENCE   121 AA;  13527 MW;  5C1D4BEC1D26B9C6 CRC64;

  Query Match             72.3%;  Score 102;  DB 1;  Length 121;
  Best Local Similarity   60.0%;  Pred. No. 1.9e-09;
  Matches   18;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | |||:|||:||:|  |||  |:|||  |:
Db     88 HAEGTYTSDISSFLRDQAAQNFVAWLKSGQ 117


RESULT 14
GLUC_LEPSP
ID   GLUC_LEPSP     STANDARD;      PRT;    78 AA.
AC   P09566;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glucagon precursor [Contains: Glucagon; Glucagon-36 (Oxyntomodulin);
DE   Glucagon-like peptide] (Fragment).
OS   Lepisosteus spatula (Alligator gar) (Atractosteus spatula).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
OC   Lepisosteus.
OX   NCBI_TaxID=7917;
RN   [1]
RP   SEQUENCE OF 1-36 AND 45-78.
RC   TISSUE=Pancreas;
RX   MEDLINE=88196798; PubMed=3282974;
RA   Pollock H.G., Kimmel J.R., Ebner K.E., Hamilton J.W., Rouse J.B.,
RA   Lance V., Rawitch A.B.;
RT   "Isolation of alligator gar (Lepisosteus spatula) glucagon,
RT   oxyntomodulin, and glucagon-like peptide: amino acid sequences of
RT   oxyntomodulin and glucagon-like peptide.";
RL   Gen. Comp. Endocrinol. 69:133-140(1988).
RN   [2]
RP   PRELIMINARY SEQUENCE OF 1-29.
RC   TISSUE=Pancreas;
RX   MEDLINE=88030594; PubMed=3311873;
RA   Pollock H.G., Kimmel J.R., Hamilton J.W., Rouse J.B., Ebner K.E.,
RA   Lance V., Rawitch A.B.;
RT   "Isolation and structures of alligator gar (Lepisosteus spatula)
RT   insulin and pancreatic polypeptide.";
RL   Gen. Comp. Endocrinol. 67:375-382(1987).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- MISCELLANEOUS: X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH
CC       AMERICAN GOOSEFISH SEQUENCES.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   PIR; S06339; GCGXA.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone.
FT   PEPTIDE       1     29       GLUCAGON.
FT   PEPTIDE       1     36       GLUCAGON-36.
FT   PEPTIDE      45     78       GLUCAGON-LIKE PEPTIDE.
SQ   SEQUENCE   78 AA;  8990 MW;  30106496271594E0 CRC64;

  Query Match             71.6%;  Score 101;  DB 1;  Length 78;
  Best Local Similarity   60.0%;  Pred. No. 1.8e-09;
  Matches   18;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | :||:||||||||  |||  |: || :|:
Db     45 HADGTYTSDVSSYLQDQAAKKFVTWLKQGQ 74


RESULT 15
GLUC_ICTPU
ID   GLUC_ICTPU     STANDARD;      PRT;    71 AA.
AC   P04093;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Glucagon precursor (Fragment).
OS   Ictalurus punctatus (Channel catfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC   Ictaluridae; Ictalurus.
OX   NCBI_TaxID=7998;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pancreas;
RX   MEDLINE=87156787; PubMed=3030323;
RA   Hoosein N.M., Mahrenholz A.M., Andrews P.C., Gurd R.S.;
RT   "Biological activities of catfish glucagon and glucagon-like
RT   peptide.";
RL   Biochem. Biophys. Res. Commun. 143:87-92(1987).
RN   [2]
RP   SEQUENCE.
RC   TISSUE=Pancreas;
RX   MEDLINE=85157536; PubMed=3838546;
RA   Andrews P.C., Ronner P.;
RT   "Isolation and structures of glucagon and glucagon-like peptide from
RT   catfish pancreas.";
RL   J. Biol. Chem. 260:3910-3914(1985).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
CC       IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- MISCELLANEOUS: X'S IN THE SEQUENCE WERE INCLUDED BY HOMOLOGY WITH
CC       AMERICAN GOOSEFISH SEQUENCES.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   PIR; A05166; GCIDC.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone.
FT   NON_TER       1      1
FT   PEPTIDE       1     29       GLUCAGON.
FT   PEPTIDE      38     71       GLUCAGON-LIKE PEPTIDE.
FT   CONFLICT     53     53       E -> D (IN REF. 2).
FT   NON_TER      71     71
SQ   SEQUENCE   71 AA;  8173 MW;  24688E79AD981A8F CRC64;

  Query Match             70.9%;  Score 100;  DB 1;  Length 71;
  Best Local Similarity   63.3%;  Pred. No. 2.4e-09;
  Matches   19;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
```

```
                | :||:||||||||  |||  || ||  |:
Db      38 HADGTYTSDVSSYLQEQAAKDFITWLKSGQ 67

Search completed: February 13, 2003, 11:02:34
Job time : 12 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 13, 2003, 11:00:06 ; Search time 29 Seconds
(without alignments)
220.257 Million cell updates/sec

Title: US-09-091-605-1
Perfect score: 141
Sequence: 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	133	94.3	180	6	Q95LG0	Q95lg0 canis famil
2	121	85.8	206	13	Q91410	Q91410 gallus gall
3	115	81.6	204	13	O12956	O12956 heloderma s
4	107	75.9	220	13	Q8UWL9	Q8uwl9 hoplobatrac
5	103	73.0	266	13	O42143	O42143 xenopus lae
6	102	72.3	72	13	Q91409	Q91409 oncorhynchu
7	102	72.3	178	13	Q91971	Q91971 oncorhynchu
8	98	69.5	219	13	O42144	O42144 xenopus lae
9	97	68.8	178	13	Q91189	Q91189 oncorhynchu
10	93	66.0	121	13	Q9DDE6	Q9dde6 brachydanio
11	89	63.1	160	13	Q9PUR1	Q9pur1 petromyzon
12	82	58.2	62	13	Q9PRW9	Q9prw9 scyliorhinu
13	78	55.3	96	13	Q9DG43	Q9dg43 ambloplites
14	70	49.6	120	13	Q9PUR0	Q9pur0 petromyzon
15	53.5	37.9	426	16	P71006	P71006 bacillus su
16	52.5	37.2	285	17	Q8TPJ9	Q8tpj9 methanosarc
17	51	36.2	810	4	Q9NTW8	Q9ntw8 homo sapien
18	51	36.2	867	4	Q9UFX9	Q9ufx9 homo sapien
19	49	34.8	130	11	Q9CVF1	Q9cvf1 mus musculu
20	49	34.8	144	11	Q9D887	Q9d887 mus musculu
21	49	34.8	236	16	Q92D11	Q92d11 listeria in
22	49	34.8	236	16	Q8Y8A1	Q8y8a1 listeria mo
23	48.5	34.4	221	16	Q9I6W5	Q9i6w5 pseudomonas
24	48	34.0	171	11	Q9D2Z7	Q9d2z7 mus musculu
25	47.5	33.7	1100	16	Q8YAE8	Q8yae8 listeria mo
26	47	33.3	302	16	Q9K3M4	Q9k3m4 streptomyce
27	46.5	33.0	328	16	Q9PM60	Q9pm60 campylobact
28	46.5	33.0	427	17	Q8TLY0	Q8tly0 methanosarc
29	46	32.6	542	10	Q9LXH0	Q91xh0 arabidopsis
30	46	32.6	815	5	Q9U1S7	Q9u1s7 caenorhabdi
31	45.5	32.3	372	10	Q9XFW9	Q9xfw9 cicer ariet
32	45	31.9	245	16	Q8X5X9	Q8x5x9 escherichia
33	45	31.9	389	2	Q93IH2	Q93ih2 wolinella s
34	45	31.9	634	3	Q9HEE5	Q9hee5 neurospora
35	45	31.9	3600	10	Q9SA64	Q9sa64 arabidopsis
36	44.5	31.6	210	5	Q95XL4	Q95xl4 caenorhabdi
37	44	31.2	136	17	Q8TLQ2	Q8tlq2 methanosarc
38	44	31.2	268	16	Q9RS44	Q9rs44 deinococcus
39	44	31.2	287	4	Q9H5K2	Q9h5k2 homo sapien
40	44	31.2	334	16	Q8U9B9	Q8u9b9 agrobacteri
41	44	31.2	387	16	Q98A97	Q98a97 rhizobium l
42	44	31.2	390	4	Q9NXC2	Q9nxc2 homo sapien
43	44	31.2	478	16	Q9KV20	Q9kv20 vibrio chol
44	44	31.2	575	9	Q38545	Q38545 bacteriopha
45	44	31.2	893	3	Q99222	Q99222 saccharomyc

ALIGNMENTS

RESULT 1
Q95LG0
ID Q95LG0 PRELIMINARY; PRT; 180 AA.
AC Q95LG0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Preproglucagon.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Irwin D.M.;
RT "cDNA cloning of proglucagon from the stomach and pancreas of the
RT dog.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF308439; AAL09425.1; -.
DR InterPro; IPR000532; Glucagon.
DR Pfam; PF00123; hormone2; 3.
DR PROSITE; PS00260; GLUCAGON; UNKNOWN_3.
SQ SEQUENCE 180 AA; 21114 MW; 80F66941AFC324FD CRC64;

Query Match 94.3%; Score 133; DB 6; Length 180;
Best Local Similarity 86.7%; Pred. No. 6.8e-14;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||| |||| |||||||||
Db      98 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 127
```

RESULT 2
Q91410
ID Q91410 PRELIMINARY; PRT; 206 AA.
AC Q91410;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Proglucagon.
GN   PROGLUCAGON.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95295739; PubMed=7776976;
RA   Irwin D.M., Wong J.;
RT   "Trout and chicken proglucagon: alternative splicing generates mRNA
RT   transcripts encoding glucagon-like peptide 2.";
RL   Mol. Endocrinol. 9:267-277(1995).
DR   EMBL; S78477; AAB34506.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 3.
SQ   SEQUENCE   206 AA;  23875 MW;  AB299E1B02FC6AA4 CRC64;

  Query Match             85.8%;  Score 121;  DB 13;  Length 206;
  Best Local Similarity   73.3%;  Pred. No. 7.5e-12;
  Matches   22;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | |||:|||::||| ||||  |||||| ||
Db    118 HAEGTYTSDITSYLEGQAAKEFIAWLVNGR 147


RESULT 3
O12956
ID   O12956      PRELIMINARY;      PRT;   204 AA.
AC   O12956; O12955;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Glucagon precursor.
OS   Heloderma suspectum (Gila monster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC   Heloderma.
OX   NCBI_TaxID=8554;
RN   [1]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC   TISSUE=INTESTINE, AND PANCREAS;
RX   MEDLINE=97172477; PubMed=9020121;
RA   Chen Y.E., Drucker D.J.;
RT   "Tissue-specific expression of unique mRNAs that encode proglucagon-
RT   derived peptides or exendin 4 in the lizard.";
RL   J. Biol. Chem. 272:4108-4115(1997).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; LPII (SHOWN HERE) AND LPI; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: ISOFORM LPII IS EXPRESSED IN BOTH PANCREAS AND
CC       INTESTINE. EXPRESSION OF ISOFORM LPI IS RESTRICTED TO THE
CC       PANCREAS. NEITHER ISOFORM IS DETECTED IN SALIVARY GLAND.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN
CC       RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   EMBL; U77612; AAB51129.1; -.
DR   EMBL; U77611; AAB51128.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Alternative splicing.
FT   SIGNAL        1     20       BY SIMILARITY.
FT   PEPTIDE      21     50       GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE     116    145       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     164    196       GLUCAGON-LIKE PEPTIDE 2.
FT   VARSPLIC    149    149       D -> E (IN ISOFORM LPI).
FT   VARSPLIC    150    204       MISSING (IN ISOFORM LPI).
SQ   SEQUENCE   204 AA;  23553 MW;  B132E3FE46873E72 CRC64;

  Query Match             81.6%;  Score 115;  DB 13;  Length 204;
  Best Local Similarity   70.0%;  Pred. No. 7.2e-11;
  Matches   21;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | :| :|||:|||| |||| ||||||| ||
Db    116 HADGRYTSDISSYLEGQAAKEFIAWLVNGR 145


RESULT 4
Q8UWL9
ID   Q8UWL9      PRELIMINARY;      PRT;   220 AA.
AC   Q8UWL9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Proglucagon.
OS   Hoplobatrachus rugulosus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae;
OC   Hoplobatrachus.
OX   NCBI_TaxID=110072;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yeung C.-M., Chow B.K.C.;
RT   "Identification of a proglucagon cDNA from Rana tigrina rugulosa that
RT   encodes two GLP-1s.";
RL   Gen. Comp. Endocrinol. 124:0-0(2001).
DR   EMBL; AF324209; AAL35758.1; -.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 4.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 4.
DR   PROSITE; PS00260; GLUCAGON; UNKNOWN_4.
SQ   SEQUENCE   220 AA;  25615 MW;  C72D926E7F89E381 CRC64;

  Query Match             75.9%;  Score 107;  DB 13;  Length 220;
  Best Local Similarity   63.3%;  Pred. No. 1.6e-09;
  Matches   19;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||::|||  :||  |: ||:|||
Db    135 HAEGTFTSDMTSYLEEKAAKEFVDWLIKGR 164


RESULT 5
O42143
ID   O42143      PRELIMINARY;      PRT;   266 AA.
AC   O42143;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Glucagon I precursor [Contains: Glucagon; glucagon-like peptide 1A
DE   (GLP-1A); glucagon-like peptide 1B (GLP-1B); glucagon-like peptide 1C
DE   (GLP-1C); glucagon-like peptide 2 (GLP-2)].
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=PANCREAS;
```

```
RX   MEDLINE=97368292; PubMed=9223287;
RA   Irwin D.M., Satkunarajah M., Wen Y., Brubaker P.L., Pederson R.A.,
RA   Wheeler M.B.;
RT   "The Xenopus proglucagon gene encodes novel GLP-1-like peptides with
RT   insulinotropic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:7915-7920(1997).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   EMBL; AF004432; AAB65660.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 5.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 5.
DR   PROSITE; PS00260; GLUCAGON; 5.
KW   Glucagon family; Hormone; Signal; Cleavage on pair of basic residues;
KW   Multigene family; Alternative splicing.
FT   SIGNAL       1      ?       POTENTIAL.
FT   PEPTIDE     53     81       GLUCAGON.
FT   PEPTIDE     97    133       GLUCAGON-LIKE PEPTIDE 1A.
FT   PEPTIDE    142    173       GLUCAGON-LIKE PEPTIDE 1B.
FT   PEPTIDE    180    211       GLUCAGON-LIKE PEPTIDE 1C.
FT   PEPTIDE    227    259       GLUCAGON-LIKE PEPTIDE 2.
FT   VARSPLIC   214    261       MISSING (IN ISOFORM 2).
SQ   SEQUENCE   266 AA;  30951 MW;  544F7BBC20AF872C CRC64;

  Query Match             73.0%;  Score 103;  DB 13;  Length 266;
  Best Local Similarity   56.7%;  Pred. No. 9.2e-09;
  Matches   17;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | |||||:|:::||  :||  |: ||:|||
Db     180 HAEGTFTNDMTNYLEEKAAKEFVGWLIKGR 209


RESULT 6
Q91409
ID   Q91409      PRELIMINARY;      PRT;    72 AA.
AC   Q91409; Q91232;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PROGLUCAGON (Fragment).
OS   Oncorhynchus tschawytscha (Chinook salmon) (King salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=74940;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95295739; PubMed=7776976;
RA   Irwin D.M., Wong J.;
RT   "Trout and chicken proglucagon: alternative splicing generates mRNA
RT   transcripts encoding glucagon-like peptide 2.";
RL   Mol. Endocrinol. 9:267-277(1995).
DR   EMBL; S78474; AAD14283.1; -.
DR   EMBL; U19920; AAC59670.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; UNKNOWN_1.
FT   NON_TER      1      1
SQ   SEQUENCE   72 AA;  8293 MW;  8584352B1C260A31 CRC64;

  Query Match             72.3%;  Score 102;  DB 13;  Length 72;
  Best Local Similarity   60.0%;  Pred. No. 3.1e-09;
  Matches   18;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | :||:|||||:||  |||  |::||  ||
Db      39 HADGTYTSDVSTYLQDQAAKDFVSWLKSGR 68


RESULT 7
Q91971
ID   Q91971      PRELIMINARY;      PRT;   178 AA.
AC   Q91971; Q91408; Q91188; Q92169;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Glucagon I precursor.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=DISTAL SMALL INTESTINE, AND PANCREAS;
RX   MEDLINE=95295739; PubMed=7776976;
RA   Irwin D.M., Wong J.;
RT   "Trout and chicken proglucagon: alternative splicing generates mRNA
RT   transcripts encoding glucagon-like peptide 2.";
RL   Mol. Endocrinol. 9:267-277(1995).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; INTESTINAL (SHOWN HERE) AND
CC       PANCREATIC; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN
CC       RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   EMBL; U19913; AAC59667.1; -.
DR   EMBL; U19917; AAC59669.1; -.
DR   EMBL; U19918; AAC60212.1; -.
DR   EMBL; U19919; AAC60213.1; -.
DR   EMBL; U19918; AAC60213.1; JOINED.
DR   EMBL; S78475; AAB34505.1; -.
DR   EMBL; S78473; AAB34504.2; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; 3.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL       1      ?       POTENTIAL.
FT   PEPTIDE      ?     49       GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT   PEPTIDE     52     80       GLUCAGON.
FT   PEPTIDE     85    120       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE    137    169       GLUCAGON-LIKE PEPTIDE 2.
FT   VARSPLIC   124    178       MISSING (IN PANCREATIC ISOFORM).
SQ   SEQUENCE   178 AA;  20034 MW;  5CF6980CF2A9D58E CRC64;

  Query Match             72.3%;  Score 102;  DB 13;  Length 178;
  Best Local Similarity   60.0%;  Pred. No. 8.6e-09;
  Matches   18;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | :||:|||||:||  |||  |::||  ||
Db      90 HADGTYTSDVSTYLQDQAAKDFVSWLKSGR 119


RESULT 8
O42144
ID   O42144      PRELIMINARY;      PRT;   219 AA.
AC   O42144;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Glucagon II precursor [Contains: Glucagon; glucagon-like peptide 1A
```

```
DE   (GLP-1A); glucagon-like peptide 1B (GLP-1B); glucagon-like peptide 1C
DE   (GLP-1C)].
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PANCREAS;
RX   MEDLINE=97368292; PubMed=9223287;
RA   Irwin D.M., Satkunarajah M., Wen Y., Brubaker P.L., Pederson R.A.,
RA   Wheeler M.B.;
RT   "The Xenopus proglucagon gene encodes novel GLP-1-like peptides with
RT   insulinotropic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:7915-7920(1997).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   EMBL; AF004433; AAB65661.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 4.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 4.
DR   PROSITE; PS00260; GLUCAGON; 3.
KW   Glucagon family; Hormone; Signal; Cleavage on pair of basic residues;
KW   Multigene family.
FT   SIGNAL        1     20       POTENTIAL.
FT   PEPTIDE      53     81       GLUCAGON.
FT   PEPTIDE      97    133       GLUCAGON-LIKE PEPTIDE 1A.
FT   PEPTIDE     142    173       GLUCAGON-LIKE PEPTIDE 1B.
FT   PEPTIDE     180    211       GLUCAGON-LIKE PEPTIDE 1C.
SQ   SEQUENCE   219 AA;  25271 MW;  ACC699233C362CE0 CRC64;

  Query Match             69.5%;  Score 98;  DB 13;  Length 219;
  Best Local Similarity   53.3%;  Pred. No. 5e-08;
  Matches   16;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||:|:::||  :||  |: ||: ||
Db     180 HAEGTFTNDMTNYLEEKAAKEFVGWLINGR 209


RESULT 9
Q91189
ID   Q91189      PRELIMINARY;      PRT;   178 AA.
AC   Q91189; Q92168;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Glucagon II precursor.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=DISTAL SMALL INTESTINE, AND PANCREAS;
RX   MEDLINE=95295739; PubMed=7776976;
RA   Irwin D.M., Wong J.;
RT   "Trout and chicken proglucagon: alternative splicing generates mRNA
RT   transcripts encoding glucagon-like peptide 2.";
RL   Mol. Endocrinol. 9:267-277(1995).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; INTESTINAL (SHOWN HERE) AND
CC       PANCREATIC; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS IN
CC       RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   EMBL; U19914; AAC59668.1; -.
DR   EMBL; U19916; AAC60210.1; -.
DR   EMBL; U19915; AAC60210.1; JOINED.
DR   EMBL; U19915; AAC60209.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 3.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 3.
DR   PROSITE; PS00260; GLUCAGON; UNKNOWN_2.
KW   Glucagon family; Hormone; Cleavage on pair of basic residues; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL        1      ?       POTENTIAL.
FT   PEPTIDE       ?     49       GRPP (GLICENTINE RELATED POLYPEPTIDE).
FT   PEPTIDE      52     80       GLUCAGON.
FT   PEPTIDE      85    120       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE     137    169       GLUCAGON-LIKE PEPTIDE 2.
FT   VARSPLIC    124    178       MISSING (IN PANCREATIC ISOFORM).
SQ   SEQUENCE   178 AA;  19998 MW;  E89D73866CD91C66 CRC64;

  Query Match             68.8%;  Score 97;  DB 13;  Length 178;
  Best Local Similarity   58.6%;  Pred. No. 5.7e-08;
  Matches   17;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKG 29
           | :||:|||||:||  |||  |::||  |
Db      90 HADGTYTSDVSTYLQDQAAKDFVSWLKSG 118


RESULT 10
Q9DDE6
ID   Q9DDE6      PRELIMINARY;      PRT;   121 AA.
AC   Q9DDE6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Glucagon polyprotein.
GN   GCG OR GLU.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99425190; PubMed=10495291;
RA   Argenton F., Zecchin E., Bortolussi M.;
RT   "Early appearance of pancreatic hormone-expressing cells in the
RT   zebrafish embryo.";
RL   Mech. Dev. 87:217-221(1999).
DR   EMBL; AJ133697; CAC20108.1; -.
DR   HSSP; P01274; 1GCN.
DR   ZFIN; ZDB-GENE-010219-1; gcg.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 1.
KW   Polyprotein.
FT   CHAIN        49     79       GLUCAGON.
FT   CHAIN        88    121       GLUCAGON-LIKE PEPTIDE 1.
SQ   SEQUENCE   121 AA;  13537 MW;  A85385F690DA180F CRC64;

  Query Match             66.0%;  Score 93;  DB 13;  Length 121;
  Best Local Similarity   63.3%;  Pred. No. 1.7e-07;
  Matches   19;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | |||:||||||||  |||  |:| |  |:
Db      88 HAEGTYTSDVSSYLQDQAAQRFVARLKSGQ 117


RESULT 11
Q9PUR1
```

```
ID   Q9PUR1      PRELIMINARY;      PRT;   160 AA.
AC   Q9PUR1; Q9PRZ8; Q9PRZ7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Glucagon I precursor [Contains: Glucagon; glucagon-like peptide 1
DE   (GLP-1); glucagon-like peptide 2 (GLP-2)].
OS   Petromyzon marinus (Sea lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Petromyzon.
OX   NCBI_TaxID=7757;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=INTESTINE;
RX   MEDLINE=20022986; PubMed=10555286;
RA   Irwin D.M., Huner O., Youson J.H.;
RT   "Lamprey proglucagon and the origin of glucagon-like peptides.";
RL   Mol. Biol. Evol. 16:1548-1557(1999).
RN   [2]
RP   SEQUENCE OF 43-71 AND 82-113.
RC   TISSUE=INTESTINE;
RX   MEDLINE=94010172; PubMed=8405897;
RA   Conlon J.M., Nielsen P.F., Youson J.H.;
RT   "Primary structures of glucagon and glucagon-like peptide isolated
RT   from the intestine of the parasitic phase lamprey Petromyzon
RT   marinus.";
RL   Gen. Comp. Endocrinol. 91:96-104(1993).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   EMBL; AF159707; AAF09186.1; -.
DR   HSSP; P01275; 1BH0.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone; Signal; Cleavage on pair of basic residues;
KW   Multigene family.
FT   SIGNAL       1     22       POTENTIAL.
FT   PEPTIDE     43     71       GLUCAGON.
FT   PEPTIDE     82    113       GLUCAGON-LIKE PEPTIDE 1.
FT   PEPTIDE    130    160       GLUCAGON-LIKE PEPTIDE 2.
SQ   SEQUENCE   160 AA;  18042 MW;  9A52C530D5A74072 CRC64;

  Query Match             63.1%;  Score 89;  DB 13;  Length 160;
  Best Local Similarity   50.0%;  Pred. No. 1.1e-06;
  Matches   14;  Conservative    8;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVK 28
          | :||||:|::|||  :||  |::|| :
Db     82 HADGTFTNDMTSYLDAKAARDFVSWLAR 109


RESULT 12
Q9PRW9
ID   Q9PRW9      PRELIMINARY;      PRT;    62 AA.
AC   Q9PRW9; Q9PRX0; Q9PRW8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Glucagon precursor [Contains: glucagon-29; glucagon-33; glucagon-like
DE   peptide] (Fragments).
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
OX   NCBI_TaxID=7830;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PANCREAS;
RX   MEDLINE=94286411; PubMed=8015974;
RA   Conlon J.M., Hazon N., Thim L.;
RT   "Primary structures of peptides derived from proglucagon isolated from
RT   the pancreas of the elasmobranch fish, Scyliorhinus canicula.";
RL   Peptides 15:163-167(1994).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone.
FT   PEPTIDE       1     29       GLUCAGON-29.
FT   PEPTIDE       1     33       GLUCAGON-33.
FT   NON_CONS     33     34
FT   PEPTIDE      34     62       GLUCAGON-LIKE PEPTIDE.
SQ   SEQUENCE   62 AA;  7270 MW;  C5FF487C12C69CD1 CRC64;

  Query Match             58.2%;  Score 82;  DB 13;  Length 62;
  Best Local Similarity   51.9%;  Pred. No. 5.3e-06;
  Matches   14;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLV 27
          | |||||| | |:  : |  |: ||:
Db      1 HSEGTFTSDYSKYMDNRRAKDFVQWLM 27


RESULT 13
Q9DG43
ID   Q9DG43      PRELIMINARY;      PRT;    96 AA.
AC   Q9DG43;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Proglucagon (Fragment).
OS   Ambloplites rupestris.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Centrarchidae; Ambloplites.
OX   NCBI_TaxID=109273;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT   "Rock Bass Proglucagon.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF190499; AAG16778.1; -.
DR   HSSP; P01274; 1GCN.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; UNKNOWN_1.
FT   NON_TER       1      1
FT   CHAIN         1    >29       GLUCAGON.
FT   CHAIN        39    >70       GLUCAGON-LIKE PEPTIDE 1.
FT   CHAIN        86    >96       GLUCAGON-LIKE PEPTIDE 2.
FT   NON_TER      96     96
SQ   SEQUENCE   96 AA;  11225 MW;  6435033EBDDC00CE CRC64;

  Query Match             55.3%;  Score 78;  DB 13;  Length 96;
  Best Local Similarity   43.3%;  Pred. No. 3.9e-05;
  Matches   13;  Conservative    7;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | :||||:| ::||  : |  || ||:  :
Db      1 HSQGTFTNDYTNYLEDRQAQDFIRWLMNNK 30


RESULT 14
Q9PUR0
ID   Q9PUR0      PRELIMINARY;      PRT;   120 AA.
AC   Q9PUR0;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Glucagon II precursor [Contains: Glucagon; glucagon-like peptide
DE   (GLP)].
OS   Petromyzon marinus (Sea lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Petromyzon.
OX   NCBI_TaxID=7757;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=INTESTINE;
RX   MEDLINE=20022986; PubMed=10555286;
RA   Irwin D.M., Huner O., Youson J.H.;
RT   "Lamprey proglucagon and the origin of glucagon-like peptides.";
RL   Mol. Biol. Evol. 16:1548-1557(1999).
CC   -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
CC       THE BLOOD SUGAR LEVEL.
CC   -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR   EMBL; AF159708; AAF09187.1; -.
DR   HSSP; P01275; 1BH0.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; hormone2; 2.
DR   PRINTS; PR00275; GLUCAGON.
DR   SMART; SM00070; GLUCA; 2.
DR   PROSITE; PS00260; GLUCAGON; 2.
KW   Glucagon family; Hormone; Signal; Cleavage on pair of basic residues;
KW   Multigene family.
FT   SIGNAL        1      ?       POTENTIAL.
FT   PEPTIDE      44     72       GLUCAGON.
FT   PEPTIDE      89    120       GLUCAGON-LIKE PEPTIDE.
SQ   SEQUENCE   120 AA;  13397 MW;  FBDE667B96E198D8 CRC64;

  Query Match             49.6%;  Score 70;  DB 13;  Length 120;
  Best Local Similarity   44.4%;  Pred. No. 0.001;
  Matches   12;  Conservative    6;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLV 27
           | :|:|||| | :|  : |  |: ||:
Db      44 HSQGSFTSDYSKHLDVKQAKDFVTWLL 70


RESULT 15
P71006
ID   P71006      PRELIMINARY;      PRT;   426 AA.
AC   P71006;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein ywhn.
GN   YWHN.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Glaser P., Sekowska A., Danchin A.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z80360; CAB02504.1; -.
DR   EMBL; Z99123; CAB15769.1; -.
DR   InterPro; IPR001431; Peptidase_M16.
DR   Pfam; PF00675; Peptidase_M16; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   426 AA;  48981 MW;  F6C43830BEAE3BF5 CRC64;

  Query Match             37.9%;  Score 53.5;  DB 16;  Length 426;
  Best Local Similarity   39.4%;  Pred. No. 2.2;
  Matches   13;  Conservative    4;  Mismatches   13;  Indels    3;  Gaps    1;

Qy       1 HXEGTFTS---DVSSYLXGQAAXXFIAWLVKGR 30
           | :|| :|   ||:||   |      ||| |:
Db     323 HSQGTHSSAYIDVASYYVTQKKEQVAAWLQYGK 355


Search completed: February 13, 2003, 11:03:11
Job time : 31 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:55:04 ; Search time 48 Seconds
(without alignments)
594.186 Million cell updates/sec

Title: US-09-091-605-2
Perfect score: 93
Sequence: 1 CATGCTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	93	100.0	528	2	US-08-835-231-8	Sequence 8, Appli
	2	93	100.0	528	4	US-09-108-661-8	Sequence 8, Appli
	3	93	100.0	955	3	US-08-784-582-57	Sequence 57, Appl
	4	93	100.0	955	3	US-08-784-582-60	Sequence 60, Appl
	5	93	100.0	2356	3	US-08-784-582-72	Sequence 72, Appl
	6	91.4	98.3	528	2	US-08-835-231-7	Sequence 7, Appli
	7	91.4	98.3	528	4	US-09-108-661-7	Sequence 7, Appli
	8	83.4	89.7	895	3	US-08-784-582-55	Sequence 55, Appl
	9	80.8	86.9	144	2	US-08-835-231-17	Sequence 17, Appl
	10	80.8	86.9	144	4	US-09-108-661-17	Sequence 17, Appl
	11	49.8	53.5	78	2	US-08-829-876-22	Sequence 22, Appl
	12	49.8	53.5	78	4	US-09-234-874A-22	Sequence 22, Appl
	13	37.8	40.6	57	2	US-08-811-028-43	Sequence 43, Appl
c	14	35.8	38.5	48	2	US-08-811-028-44	Sequence 44, Appl
c	15	32.8	35.3	53	2	US-08-811-028-46	Sequence 46, Appl
c	16	32.2	34.6	73	2	US-08-829-876-24	Sequence 24, Appl
c	17	32.2	34.6	73	4	US-09-234-874A-24	Sequence 24, Appl
c	18	29.4	31.6	52	2	US-08-829-876-23	Sequence 23, Appl
c	19	29.4	31.6	52	4	US-09-234-874A-23	Sequence 23, Appl
	20	28	30.1	2409	1	US-08-392-828C-1	Sequence 1, Appli
	21	28	30.1	2409	3	US-09-330-945-1	Sequence 1, Appli
	22	26.2	28.2	36	2	US-08-811-028-45	Sequence 45, Appl
	23	26.2	28.2	5558	4	US-08-961-527-103	Sequence 103, App
c	24	25.8	27.7	144	2	US-08-835-231-17	Sequence 17, Appl
c	25	25.8	27.7	144	4	US-09-108-661-17	Sequence 17, Appl
	26	25.4	27.3	1614	4	US-09-522-217-7	Sequence 7, Appli
	27	25.4	27.3	2665	3	US-09-040-005-1	Sequence 1, Appli
c	28	25	26.9	725	4	US-09-221-017B-198	Sequence 198, App
	29	24.6	26.5	516	1	US-08-532-828B-13	Sequence 13, Appl
	30	24.6	26.5	1263	1	US-08-532-828B-11	Sequence 11, Appl
	31	24.6	26.5	1643	1	US-08-532-828B-1	Sequence 1, Appli
	32	24.6	26.5	1643	1	US-08-532-828B-7	Sequence 7, Appli
	33	24.6	26.5	1643	1	US-08-532-828B-9	Sequence 9, Appli
	34	24.6	26.5	1643	1	US-08-700-359-7	Sequence 7, Appli
	35	24.6	26.5	1643	1	US-08-700-359-8	Sequence 8, Appli
	36	24.6	26.5	1643	1	US-08-700-359-10	Sequence 10, Appl
	37	24.6	26.5	1643	1	US-08-674-168-21	Sequence 21, Appl
	38	24.6	26.5	1643	2	US-08-596-366-5	Sequence 5, Appli
	39	24.6	26.5	1643	2	US-08-596-366-7	Sequence 7, Appli
	40	24.6	26.5	1643	2	US-08-967-104-5	Sequence 5, Appli
	41	24.6	26.5	1643	2	US-08-967-104-7	Sequence 7, Appli
	42	24.6	26.5	1643	3	US-08-985-908-3	Sequence 3, Appli
	43	24.6	26.5	1643	3	US-08-985-908-4	Sequence 4, Appli
	44	24.6	26.5	1643	3	US-08-985-908-6	Sequence 6, Appli
	45	24.6	26.5	1643	3	US-08-852-730-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-835-231-8
; Sequence 8, Application US/08835231
; Patent No. 5861284
; GENERAL INFORMATION:
; APPLICANT: NISHIMURA, Osamu
; APPLICANT: KURIYAMA, Masato
; APPLICANT: KOYAMA, No. 5861284uyuki
; APPLICANT: FUKUDA, Tsunehiko
; TITLE OF INVENTION: METHOD FOR PRODUCING A BIOLOGICALLY
; TITLE OF INVENTION: ACTIVE RECOMBINANT CYSTEINE-FREE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,231
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,709
; FILING DATE: 07-DEC-1994
; APPLICATION NUMBER: 07/838,857
; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: JP 024841
; FILING DATE: 19-FEB-1991
; APPLICATION NUMBER: JP 0271438
; FILING DATE: 18-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DAVID, RESNICK S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 41614-FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: 200291 STRE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-08-835-231-8

  Query Match             100.0%;  Score 93;  DB 2;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-26;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 2
US-09-108-661-8
; Sequence 8, Application US/09108661
; Patent No. 6287806
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 6287806uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/108,661
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  528 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-09-108-661-8

  Query Match             100.0%;  Score 93;  DB 4;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-26;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 3
US-08-784-582-57
; Sequence 57, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  955 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-784-582-57
```

```
  Query Match             100.0%;  Score 93;  DB 3;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-26;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 4
US-08-784-582-60
; Sequence 60, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  955 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-784-582-60

  Query Match             100.0%;  Score 93;  DB 3;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-26;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 5
US-08-784-582-72
; Sequence 72, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2356 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-784-582-72

  Query Match             100.0%;  Score 93;  DB 3;  Length 2356;
  Best Local Similarity   100.0%;  Pred. No. 1e-25;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1704 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 1763

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    1764 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 1796
```

```
RESULT 6
US-08-835-231-7
; Sequence 7, Application US/08835231
; Patent No. 5861284
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 5861284uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/835,231
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  528 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-08-835-231-7

  Query Match             98.3%;  Score 91.4;  DB 2;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 2.4e-25;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACGAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 7
```

```
US-09-108-661-7
; Sequence 7, Application US/09108661
; Patent No. 6287806
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 6287806uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/108,661
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  528 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-09-108-661-7

  Query Match             98.3%;  Score 91.4;  DB 4;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 2.4e-25;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACGAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 8
US-08-784-582-55
```

```
; Sequence 55, Application US/08784582
; Patent No. 6110707
; GENERAL INFORMATION:
;   APPLICANT:  Newgard, Christopher B.
;   APPLICANT:  Halban, Philippe A.
;   APPLICANT:  No. 6110707mington, Karl D.
;   APPLICANT:  Clark, Samuel A.
;   APPLICANT:  Thigpen, Anice E.
;   APPLICANT:  Quaade, Christian
;   APPLICANT:  Kruse, Fred
;   APPLICANT:  McGarry, Dennis
;   TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;   TITLE OF INVENTION:  SECRETORY CELL LINES
;   NUMBER OF SEQUENCES:  79
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Arnold, White & Durkee
;     STREET:  P.O. Box 4433
;     CITY:  Houston
;     STATE:  Texas
;     COUNTRY:  USA
;     ZIP:  77210
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/784,582
;     FILING DATE:  Concurrently Herewith
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 60/028,427
;     FILING DATE:  15-OCT-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/589,028
;     FILING DATE:  19-JAN-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Highlander, Steven L.
;     REGISTRATION NUMBER:  37,642
;     REFERENCE/DOCKET NUMBER:  UTSD:514
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  512/418-3000
;     TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH:  895 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
US-08-784-582-55

  Query Match             89.7%;  Score 83.4;  DB 3;  Length 895;
  Best Local Similarity   93.5%;  Pred. No. 3e-22;
  Matches   87;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     343 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 402

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     403 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 435


RESULT 9
US-08-835-231-17
; Sequence 17, Application US/08835231
; Patent No. 5861284
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 5861284uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/835,231
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  144 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  22...144
US-08-835-231-17

  Query Match             86.9%;  Score 80.8;  DB 2;  Length 144;
  Best Local Similarity   92.4%;  Pred. No. 1.5e-21;
  Matches   85;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||| |||||||||| |||||||| |||||| |||||||||| ||||||||
Db      43 CATGCTGAAGGCACCTTTACCAGCGATGTAAGCTCTTATCTGGAAGGCCAGGCTGCCAAA 102

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGG 92
           |||||||||||||||||||||||||||||| ||
Db     103 GAATTCATTGCTTGGCTGGTGAAAGGCCGTGG 134


RESULT 10
US-09-108-661-17
; Sequence 17, Application US/09108661
; Patent No. 6287806
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 6287806uyuki
```

```
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/108,661
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  144 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  22...144
US-09-108-661-17

  Query Match             86.9%;  Score 80.8;  DB 4;  Length 144;
  Best Local Similarity   92.4%;  Pred. No. 1.5e-21;
  Matches   85;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||| |||||||||| |||||||| |||||| |||||||||| ||||||||
Db      43 CATGCTGAAGGCACCTTTACCAGCGATGTAAGCTCTTATCTGGAAGGCCAGGCTGCCAAA 102

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGG 92
           |||||||||||||||||||||||||||||| ||
Db     103 GAATTCATTGCTTGGCTGGTGAAAGGCCGTGG 134


RESULT 11
US-08-829-876-22
; Sequence 22, Application US/08829876
; Patent No. 5962266
;  GENERAL INFORMATION:
;    APPLICANT:  White, Tyler R.
;    APPLICANT:  Damm, Deborah
;    APPLICANT:  Lesikar, David D.
;    APPLICANT:  McFadden, Kathleen
;    APPLICANT:  Garrick, Brett L.
;    TITLE OF INVENTION:  PROTEASE INHIBITOR PEPTIDES
;    NUMBER OF SEQUENCES:  228
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/829,876
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/436,555
;      FILING DATE:  08-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pelto, Don J.
;      REGISTRATION NUMBER:  33,754
;      REFERENCE/DOCKET NUMBER:  56324/106/SCNO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-829-876-22

  Query Match             53.5%;  Score 49.8;  DB 2;  Length 78;
  Best Local Similarity   82.6%;  Pred. No. 5.9e-10;
  Matches   57;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      24 TGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAA 83
           ||| ||   |||||| |||||||| |||||||| |||||||||| |||||| |||| ||
Db       1 TGACGTCTCTTCTTACTTGGAAGGTCAAGCTGCTAAGGAATTCATCGCTTGGTTGGTCAA 60

Qy      84 AGGCCGAGG 92
           ||| ||||
Db      61 AGGTAGAGG 69


RESULT 12
US-09-234-874A-22
; Sequence 22, Application US/09234874A
; Patent No. 6376648
;  GENERAL INFORMATION:
;       APPLICANT: White, Tyler R.
;                  Damm, Deborah
;                  Lesikar, David D.
;                  McFadden, Kathleen
;                  Garrick, Brett L.
;       TITLE OF INVENTION: PROTEASE INHIBITOR PEPTIDES
;       NUMBER OF SEQUENCES: 228
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Foley & Lardner
;            STREET: 3000 K Street, N.W., Suite 500
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: USA
;            ZIP: 20007-5109
```

```
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/234,874A
;              FILING DATE: 11-Jun-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 08/436,555
;              FILING DATE: 08-MAY-1995
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Bent, Stephen
;              REGISTRATION NUMBER: 29,768
;              REFERENCE/DOCKET NUMBER: 056324/0106
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (202)672-5300
;              TELEFAX: (202)672-5399
;              TELEX: 904136
;    INFORMATION FOR SEQ ID NO: 22:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 78 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;          MOLECULE TYPE: DNA (genomic)
;          SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-234-874A-22

  Query Match             53.5%;  Score 49.8;  DB 4;  Length 78;
  Best Local Similarity   82.6%;  Pred. No. 5.9e-10;
  Matches   57;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      24 TGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAA 83
           ||| ||   |||||| |||||||| |||||||| |||||||||| |||||| |||| ||
Db       1 TGACGTCTCTTCTTACTTGGAAGGTCAAGCTGCTAAGGAATTCATCGCTTGGTTGGTCAA 60

Qy      84 AGGCCGAGG 92
           ||| ||||
Db      61 AGGTAGAGG 69


RESULT 13
US-08-811-028-43
; Sequence 43, Application US/08811028C
; Patent No. 5891671
; GENERAL INFORMATION:
;  APPLICANT: SUZUKI, Yuji
;  APPLICANT: MAGOTA, Koji
;  APPLICANT: MASUDA, Toyofumi
;  TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING
;  TITLE OF INVENTION: ENZYME
;  FILE REFERENCE: 001560-294
;  CURRENT APPLICATION NUMBER: US/08/811,028C
;  CURRENT FILING DATE: 1987-03-04
;  EARLIER APPLICATION NUMBER: JP 8-70906
;  EARLIER FILING DATE: 1996-03-04
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
;   OTHER INFORMATION: GLP-1
US-08-811-028-43

  Query Match             40.6%;  Score 37.8;  DB 2;  Length 57;
  Best Local Similarity   78.9%;  Pred. No. 1.8e-05;
  Matches   45;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       5 CTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGG 61
           | ||||| |||||||||| ||||| || || ||| ||||||| || || || || |
Db       1 CGGAAGGTACCTTTACCAGCGATGTGAGCTCGTATCTGGAAGGTCAGGCGGCAAAAG 57


RESULT 14
US-08-811-028-44/c
; Sequence 44, Application US/08811028C
; Patent No. 5891671
; GENERAL INFORMATION:
;  APPLICANT: SUZUKI, Yuji
;  APPLICANT: MAGOTA, Koji
;  APPLICANT: MASUDA, Toyofumi
;  TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING
;  TITLE OF INVENTION: ENZYME
;  FILE REFERENCE: 001560-294
;  CURRENT APPLICATION NUMBER: US/08/811,028C
;  CURRENT FILING DATE: 1987-03-04
;  EARLIER APPLICATION NUMBER: JP 8-70906
;  EARLIER FILING DATE: 1996-03-04
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
;   OTHER INFORMATION: GLP-2
US-08-811-028-44

  Query Match             38.5%;  Score 35.8;  DB 2;  Length 48;
  Best Local Similarity   85.1%;  Pred. No. 9.4e-05;
  Matches   40;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGG 47
           ||||| ||||| |||||||||||| ||||| || || ||| |||||||
Db      48 CATGCGGAAGGTACCTTTACCAGCGATGTGAGCTCGTATCTGGAAGG 2


RESULT 15
US-08-811-028-46/c
; Sequence 46, Application US/08811028C
; Patent No. 5891671
; GENERAL INFORMATION:
;  APPLICANT: SUZUKI, Yuji
;  APPLICANT: MAGOTA, Koji
;  APPLICANT: MASUDA, Toyofumi
;  TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING
;  TITLE OF INVENTION: ENZYME
;  FILE REFERENCE: 001560-294
;  CURRENT APPLICATION NUMBER: US/08/811,028C
;  CURRENT FILING DATE: 1987-03-04
;  EARLIER APPLICATION NUMBER: JP 8-70906
;  EARLIER FILING DATE: 1996-03-04
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
;   OTHER INFORMATION: GLP-4
US-08-811-028-46

  Query Match             35.3%;  Score 32.8;  DB 2;  Length 53;
  Best Local Similarity   84.1%;  Pred. No. 0.0013;
  Matches   37;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      49 CAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGG 92
           || || || || |||||||| || ||||||||||||||||| ||
Db      53 CAGGCGGCAAAAGAATTCATCGCGTGGCTGGTGAAAGGCCGTGG 10
```

Search completed: February 14, 2003, 08:25:50
Job time : 50 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 13, 2003, 11:02:22 ; Search time 15 Seconds
(without alignments)
60.807 Million cell updates/sec

Title: US-09-091-605-1
Perfect score: 141
Sequence: 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	133	94.3	30	1	US-08-066-480-6	Sequence 6, Appli
2	133	94.3	30	1	US-08-095-162-1	Sequence 1, Appli
3	133	94.3	30	1	US-08-470-220A-1	Sequence 1, Appli
4	133	94.3	30	2	US-08-927-227-1	Sequence 1, Appli
5	133	94.3	30	3	US-08-967-374-1	Sequence 1, Appli
6	133	94.3	30	4	US-09-348-136-1	Sequence 5, Appli
7	133	94.3	30	4	US-08-961-405A-5	Sequence 5, Appli
8	133	94.3	30	4	US-08-915-918A-5	Sequence 4, Appli
9	133	94.3	30	4	US-09-302-596-4	Sequence 3, Appli
10	133	94.3	30	4	US-08-472-349-3	Sequence 4, Appli
11	133	94.3	30	4	US-09-333-415-4	Sequence 4, Appli
12	133	94.3	30	4	US-09-585-181A-4	Sequence 10, Appl
13	133	94.3	30	4	US-09-209-799D-10	Sequence 1, Appli
14	133	94.3	30	4	US-09-975-905-1	Sequence 1, Appli
15	133	94.3	30	4	US-09-505-991-1	Sequence 1, Appli
16	133	94.3	30	4	US-09-573-809-1	Sequence 4, Appli
17	133	94.3	30	4	US-09-303-016-4	Sequence 4, Appli
18	133	94.3	30	4	US-09-212-663-4	Sequence 27, Appl
19	133	94.3	30	5	PCT-US95-15800-27	Sequence 1, Appli
20	133	94.3	31	1	US-09-025-951-1	Sequence 3, Appli
21	133	94.3	31	1	US-08-095-162-3	Sequence 1, Appli
22	133	94.3	31	1	US-08-295-913A-1	Sequence 3, Appli
23	133	94.3	31	1	US-08-470-220A-3	Sequence 3, Appli
24	133	94.3	31	2	US-08-807-263-3	Sequence 3, Appli
25	133	94.3	31	3	US-08-967-374-3	Sequence 1, Appli
26	133	94.3	31	4	US-08-961-405A-1	Sequence 1, Appli
27	133	94.3	31	4	US-09-258-750-1	Sequence 1, Appli
28	133	94.3	31	4	US-08-915-918A-1	Sequence 1, Appli
29	133	94.3	31	4	US-09-302-596-3	Sequence 3, Appli
30	133	94.3	31	4	US-08-472-349-2	Sequence 2, Appli
31	133	94.3	31	4	US-09-623-618B-2	Sequence 2, Appli
32	133	94.3	31	4	US-09-623-618B-17	Sequence 17, Appl
33	133	94.3	31	4	US-09-623-618B-27	Sequence 27, Appl
34	133	94.3	31	4	US-09-623-618B-28	Sequence 28, Appl
35	133	94.3	31	4	US-09-333-415-3	Sequence 3, Appli
36	133	94.3	31	4	US-09-209-799D-1	Sequence 1, Appli
37	133	94.3	31	4	US-09-209-799D-29	Sequence 29, Appl
38	133	94.3	31	4	US-09-265-141A-1	Sequence 1, Appli
39	133	94.3	31	4	US-09-505-991-3	Sequence 3, Appli
40	133	94.3	31	4	US-09-303-016-3	Sequence 3, Appli
41	133	94.3	31	4	US-09-398-111-1	Sequence 1, Appli
42	133	94.3	31	4	US-09-212-663-3	Sequence 3, Appli
43	133	94.3	31	5	PCT-US95-15800-28	Sequence 28, Appl
44	133	94.3	32	4	US-09-258-750-85	Sequence 85, Appl
45	133	94.3	32	4	US-09-398-111-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-08-066-480-6
; Sequence 6, Application US/08066480
; Patent No. 5424286
; GENERAL INFORMATION:
; APPLICANT: Eng, John
; TITLE OF INVENTION: Pharmaceutical Compositions And Use of
; TITLE OF INVENTION: Exendin-3 and Exendin-4 for Treatment of Diabetes Mellitus
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/066,480
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 93,084
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..30
; OTHER INFORMATION: /label= GLP-1-7-36
; OTHER INFORMATION: /note= "GLP-1(7-36) fragment"
US-08-066-480-6

Query Match 94.3%; Score 133; DB 1; Length 30;
Best Local Similarity 86.7%; Pred. No. 1.3e-14;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 2
US-08-095-162-1
; Sequence 1, Application US/08095162
; Patent No. 5512459
;  GENERAL INFORMATION:
;    APPLICANT:  Wagner, Fred W.
;    APPLICANT:  Stout, Jay
;    APPLICANT:  Henriksen, Dennis
;    APPLICANT:  Partridge, Bruce
;    APPLICANT:  Manning, Shane
;    TITLE OF INVENTION:  Enzymatic Method for Modification of
;    TITLE OF INVENTION:  Recombinant Polypeptides
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  3100 No. 5512459west Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/095,162
;      FILING DATE:  20-JUL-1993
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nelson, Albin J.
;      REGISTRATION NUMBER:  28,659
;      REFERENCE/DOCKET NUMBER:  8648.32-US01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-332-5300
;      TELEFAX:  612-332-9081
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      CLONE:  GLP1 7-36-NH2  (Glucagon-like Peptide)
US-08-095-162-1

  Query Match             94.3%;  Score 133;  DB 1;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 3
US-08-470-220A-1
; Sequence 1, Application US/08470220A
; Patent No. 5707826
;  GENERAL INFORMATION:
;    APPLICANT:  Wagner, Fred W.
;    APPLICANT:  Stout, Jay
;    APPLICANT:  Henriksen, Dennis
;    APPLICANT:  Partridge, Bruce
;    APPLICANT:  Manning, Shane
;    TITLE OF INVENTION:  Enzymatic Method for Modification of
;    TITLE OF INVENTION:  Recombinant Polypeptides
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  3100 No. 5707826west Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,220A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/095,162
;      FILING DATE:  20-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nelson, Albin J.
;      REGISTRATION NUMBER:  28,659
;      REFERENCE/DOCKET NUMBER:  8648.32-US01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-332-5300
;      TELEFAX:  612-332-9081
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      CLONE:  GLP1 7-36-NH2  (Glucagon-like Peptide)
US-08-470-220A-1

  Query Match             94.3%;  Score 133;  DB 1;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 4
US-08-927-227-1
; Sequence 1, Application US/08927227A
; Patent No. 5977071
; GENERAL INFORMATION:
;  APPLICANT: Galloway, James A.
;  APPLICANT: Hoffmann, James A.
;  TITLE OF INVENTION: GLUCAGON-LIKE INSULINOTROPIC PEPTIDE ANALGOGS,
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS
;  FILE REFERENCE: X-9332B
;  CURRENT APPLICATION NUMBER: US/08/927,227A
;  CURRENT FILING DATE: 1997-09-10
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: The arginine residue at position 30 is modified so
;   OTHER INFORMATION: as to replace the terminal carboxyl group with an
;   OTHER INFORMATION: amine.
US-08-927-227-1

  Query Match             94.3%;  Score 133;  DB 2;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 5
US-08-967-374-1
; Sequence 1, Application US/08967374
; Patent No. 6037143
;  GENERAL INFORMATION:
;    APPLICANT:  Wagner, Fred W.
;    APPLICANT:  Stout, Jay
;    APPLICANT:  Henriksen, Dennis
;    APPLICANT:  Partridge, Bruce
;    APPLICANT:  Manning, Shane
;    TITLE OF INVENTION:  Enzymatic Method for Modification of
;    TITLE OF INVENTION:  Recombinant Polypeptides
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  3100 No. 6037143west Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,374
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/520,485
;      FILING DATE:  29-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Carter, Charles G.
;      REGISTRATION NUMBER:  35,093
;      REFERENCE/DOCKET NUMBER:  8648.32-USD1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-332-5300
;      TELEFAX:  612-332-9081
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      CLONE:  GLP1 7-36-NH2  (Glucagon-like Peptide)
US-08-967-374-1

  Query Match             94.3%;  Score 133;  DB 3;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 6
US-09-348-136-1
; Sequence 1, Application US/09348136
; Patent No. 6133235
; GENERAL INFORMATION:
;  APPLICANT:  Galloway, James A.
;  APPLICANT:  Hoffmann, James A.
;  TITLE OF INVENTION:  GLUCAGON-LIKE INSULINOTROPIC PEPTIDE ANALGOGS,
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS
;  FILE REFERENCE:  X-9332B
;  CURRENT APPLICATION NUMBER: US/09/348,136
;  CURRENT FILING DATE:  1999-07-06
;  PRIOR APPLICATION NUMBER:  US 08/927,227
;  PRIOR FILING DATE:  1997-09-10
;  NUMBER OF SEQ ID NOS:  1
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: The arginine residue at position 30 is modified so
;   OTHER INFORMATION: as to replace the terminal carboxyl group with an
;   OTHER INFORMATION: amine.
US-09-348-136-1

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 7
US-08-961-405A-5
; Sequence 5, Application US/08961405A
; Patent No. 6191102
;  GENERAL INFORMATION:
;    APPLICANT:  DiMarchi, Richard D.
;    APPLICANT:  Efendic, Suad
;    TITLE OF INVENTION:  USE OF GLP-1 ANALOGS AND DERIVATIVES
;    TITLE OF INVENTION:  ADMINISTERED PERIPHERALLY IN REGULATION OF OBESITY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BARNES & THORNBURG
;      STREET:  200 W. Madison, Suite 2601
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,405A
;      FILING DATE:  30-OCT-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/030,213
;      FILING DATE:  05-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  3051/90264
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-357-1313
;      TELEFAX:  312-759-5646
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-961-405A-5

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 8
US-08-915-918A-5
; Sequence 5, Application US/08915918A
; Patent No. 6277819
;  GENERAL INFORMATION:
;    APPLICANT:  Efendic, Suad
;    TITLE OF INVENTION:  USE OF GLP-1 OR ANALOGS IN TREATMENT OF
;    TITLE OF INVENTION:  MYOCARDIAL INFARCTION
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,918A
;      FILING DATE:  21-AUG-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  8792/28
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-915-918A-5

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 9
US-09-302-596-4
; Sequence 4, Application US/09302596
; Patent No. 6284725
; GENERAL INFORMATION:
;  APPLICANT: Coolidge, Thomas R.
;  APPLICANT:  Ehlers, Mario R.W.
;  TITLE OF INVENTION: Metabolic Intervention with GLP-1 to Improve the Function of
;  TITLE OF INVENTION:  Ischemic and Reperfused Tissue
;  FILE REFERENCE: P03660US1
;  CURRENT APPLICATION NUMBER: US/09/302,596
;  CURRENT FILING DATE:  1999-04-30
;  PRIOR APPLICATION NUMBER: 60/103,498
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: mammalian
US-09-302-596-4

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 10
US-08-472-349-3
; Sequence 3, Application US/08472349
; Patent No. 6284727
;  GENERAL INFORMATION:
;    APPLICANT:  Kim, Yesook
;    APPLICANT:  Lambert, William J.
;    APPLICANT:  Qi, Hong
;    APPLICANT:  Gelfand, Robert A.
;    APPLICANT:  Geoghegan, Kieran F.
;    APPLICANT:  Danley, Dennis E.
;    TITLE OF INVENTION:  Prolonged Delivery of Peptides
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pfizer Inc
;      STREET:  235 East 42nd Street, 20th Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10017-5755
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,349
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/181,655
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sheyka, Robert F.
;      REGISTRATION NUMBER:  31,304
;      REFERENCE/DOCKET NUMBER:  PC8391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)573-1189
;      TELEFAX:  (212)573-1939
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;      STRAIN:  N/A
;      INDIVIDUAL ISOLATE:  N/A
;      HAPLOTYPE:  N/A
;      CELL LINE:  N/A
;    IMMEDIATE SOURCE:
;      LIBRARY:  N/A
;      CLONE:  N/A
```

```
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  N/A
;      MAP POSITION:  N/A
;      UNITS:  N/A
US-08-472-349-3

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 11
US-09-333-415-4
; Sequence 4, Application US/09333415
; Patent No. 6344180
; GENERAL INFORMATION:
;  APPLICANT:  Holst, Jens J.
;  APPLICANT:  Vilsboll, Tina
;  TITLE OF INVENTION:  GLP-1 as a Diagnostic Test to Determine Beta-Cell
;  TITLE OF INVENTION:  Function and the Presence of the Condition of IGT and
;  TITLE OF INVENTION:  Type-II Diabetes
;  FILE REFERENCE:  P03987US0
;  CURRENT APPLICATION NUMBER: US/09/333,415
;  CURRENT FILING DATE:  1999-06-15
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 4
;    LENGTH: 30
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-333-415-4

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 12
US-09-585-181A-4
; Sequence 4, Application US/09585181A
; Patent No. 6358924
; GENERAL INFORMATION:
;  APPLICANT: Hoffmann, James
;  TITLE OF INVENTION: GLP-1 FORMULATIONS
;  FILE REFERENCE: X-11368
;  CURRENT APPLICATION NUMBER: US/09/585,181A
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/067,600
;  PRIOR FILING DATE: 1997-12-05
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;    LENGTH: 30
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: MOD_RES
;    LOCATION: (30)..(30)
;    OTHER INFORMATION: AMIDATION
US-09-585-181A-4

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 13
US-09-209-799D-10
; Sequence 10, Application US/09209799D
; Patent No. 6380357
; GENERAL INFORMATION:
;  APPLICANT: Hermeling, Ronald
;  APPLICANT:  Hoffmann, James
;  APPLICANT:  Narasimhan, Chakravarthy
;  TITLE OF INVENTION: GLUCAGON-LIKE PEPTIDE-1 CRYSTALS
;  FILE REFERENCE: X-10242
;  CURRENT APPLICATION NUMBER: US/09/209,799D
;  CURRENT FILING DATE:  1998-12-11
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;    LENGTH: 30
;    TYPE: PRT
;    ORGANISM: Artificial
;    FEATURE:
;    OTHER INFORMATION: synthetic construct
US-09-209-799D-10

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 14
US-09-975-905-1
; Sequence 1, Application US/09975905
; Patent No. 6388053
; GENERAL INFORMATION:
;  APPLICANT: Galloway, John  A
;  APPLICANT:  Hoffmann, James A
;  TITLE OF INVENTION: Glucagon-Like Insulinotropic Peptides, Compositions and Method
;  FILE REFERENCE: X-9332E
;  CURRENT APPLICATION NUMBER: US/09/975,905
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 09/573,809
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;    LENGTH: 30
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: MOD_RES
;    LOCATION: (30)..(30)
;    OTHER INFORMATION: The arginine residue at position 30 is modified so as to repla
;    OTHER INFORMATION: the terminal carboxyl group with an amine.
US-09-975-905-1

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 15
US-09-505-991-1
```

```
; Sequence 1, Application US/09505991
; Patent No. 6403361
;  GENERAL INFORMATION:
;       APPLICANT: Wagner, Fred W.
;                  Stout, Jay
;                  Henriksen, Dennis
;                  Partridge, Bruce
;                  Manning, Shane
;       TITLE OF INVENTION: Enzymatic Method for Modification of
;                           Recombinant Polypeptides
;       NUMBER OF SEQUENCES: 26
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Merchant & Gould
;            STREET: 3100 No. 6403361west Center
;            CITY: Minneapolis
;            STATE: MN
;            COUNTRY: USA
;            ZIP: 55402
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/505,991
;            FILING DATE: 17-Feb-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/520,485
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Carter, Charles G.
;            REGISTRATION NUMBER: 35,093
;            REFERENCE/DOCKET NUMBER: 8648.32-USD1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 612-332-5300
;            TELEFAX: 612-332-9081
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 30 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       IMMEDIATE SOURCE:
;            CLONE: GLP1 7-36-NH2  (Glucagon-like Peptide)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-505-991-1

  Query Match             94.3%;  Score 133;  DB 4;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 1.3e-14;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||  ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


Search completed: February 13, 2003, 11:03:53
Job time : 16 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: February 13, 2003, 10:57:11 ; Search time 35 Seconds
(without alignments)
118.022 Million cell updates/sec

Title: US-09-091-605-1
Perfect score: 141
Sequence: 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGRX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	134	95.0	31	17	AAW03901	Glucagon like pept
2	134	95.0	31	17	AAW03902	Glucagon like pept
3	133	94.3	30	15	AAR45435	Insulinotropin der
4	133	94.3	30	15	AAR63247	Insulinotropin (GL
5	133	94.3	30	16	AAR69063	Amidated Glucagon
6	133	94.3	30	16	AAR79809	Glucagon like pept
7	133	94.3	30	16	AAR80548	Human glucagon lik
8	133	94.3	30	17	AAR98956	Target peptide (GL
9	133	94.3	30	17	AAR98975	GLP1(7-35)-NH2. S
10	133	94.3	30	18	AAW16383	Glucagon-like pept
11	133	94.3	30	19	AAW29555	Thr-8 GLP-1(7-36)-
12	133	94.3	30	19	AAW63288	Glucagon-like pept
13	133	94.3	30	19	AAW63182	GLP-1(7-36). Homo
14	133	94.3	30	19	AAW50906	Glucagon-like pept
15	133	94.3	30	20	AAY42935	Glucagon-like pept
16	133	94.3	30	20	AAY42937	Arg26-GLP-1(7-36).
17	133	94.3	30	20	AAY27374	Glucagon-like pept
18	133	94.3	30	20	AAY39773	Glucagon like pept
19	133	94.3	30	20	AAY34198	GLP-1 mutant pepti
20	133	94.3	30	20	AAY31503	Glucagon-like pept
21	133	94.3	30	20	AAY22166	GLP-1-like peptide
22	133	94.3	30	20	AAY03719	Amino acid sequenc
23	133	94.3	30	21	AAB11283	GLP-1 peptide SEQ
24	133	94.3	30	21	AAB21336	GLP-1 analogue Thr
25	133	94.3	30	21	AAB21340	GLP-1 peptide GLP-
26	133	94.3	30	21	AAB21352	GLP-1 analogue #12
27	133	94.3	30	21	AAB21108	Human glucagon-lik
28	133	94.3	30	21	AAB07294	Modified Glucagon
29	133	94.3	30	21	AAB07313	Modified Glucagon
30	133	94.3	30	21	AAB07314	Modified Glucagon
31	133	94.3	30	21	AAY53280	Glucagon-like pept
32	133	94.3	30	21	AAY78949	Glucagon-like pept
33	133	94.3	30	22	AAU07375	Mammalian glucagon
34	133	94.3	30	22	AAE09260	Human glucagon-lik
35	133	94.3	30	22	AAG63287	An insoluble gluca
36	133	94.3	30	22	AAG63289	An insoluble gluca
37	133	94.3	30	22	AAG63303	An insoluble gluca
38	133	94.3	30	22	AAB82336	Glucagon-like pept
39	133	94.3	30	22	AAB83291	GLP-1 peptide #2.
40	133	94.3	30	22	AAG70461	GLP-1. Unidentifi
41	133	94.3	30	22	AAB91170	Pancreatic hormone
42	133	94.3	30	22	AAB91181	Pancreatic hormone
43	133	94.3	30	22	AAB60124	Human glucagon-lik
44	133	94.3	30	22	AAB60127	Human glucagon-lik
45	133	94.3	30	22	AAB60249	Glucagon-like pept

ALIGNMENTS

RESULT 1
AAW03901
ID AAW03901 standard; peptide; 31 AA.
XX
AC AAW03901;
XX
DT 15-APR-1997 (first entry)
XX
DE Glucagon like peptide 1 (7-37) analogue Ser26.
XX
KW Human; glucagon like peptide; GLP-1; analogue; stimulation;
KW pancreas; insulin; islet cell; treatment; type II diabetes.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 20
FT /note= "wild type Lys substituted with Ser"
FT Misc-difference 29
FT /note= "optionally absent when Arg30 and Gly31 are
FT absent"
FT Misc-difference 30
FT /note= "optionally absent when Gly31 is absent"
FT Misc-difference 31
FT /note= "optionally absent"
XX
PN US5545618-A.
XX
PD 13-AUG-1996.
XX
PF 24-JAN-1990; 90US-0468736.
XX
PR 20-SEP-1991; 91US-0762768.

```
PR   24-JAN-1990;   90US-0468736.
PR   10-DEC-1993;   93US-0165516.
XX
PA   (BUCK/) BUCKLEY D I.
PA   (HABE/) HABENER J F.
PA   (MALL/) MALLORY J B.
PA   (MOJS/) MOJSOV S.
XX
PI   Buckley DI,  Habener JF,  Mallory JB,  Mojsov S;
XX
DR   WPI; 1996-383697/38.
XX
PT   New modified glucagon-like peptide I fragments - have higher
PT   activity than glucagon or have improved plasma stability, useful for
PT   treating type II diabetes
XX
PS   Example 1; page -; 16pp; English.
XX
CC   The present peptide is a specific example of a claimed human
CC   glucagon like peptide 1 (GLP-1) analogue, which is useful for
CC   stimulating insulin release from pancreatic islet cells, especially
CC   in the treatment of type II diabetes at doses of 1 pg/kg to
CC   1 mg/kg.
XX
SQ   Sequence   31 AA;

  Query Match             95.0%;  Score 134;  DB 17;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.7e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAASEFIAWLVKGR 30


RESULT 2
AAW03902
ID   AAW03902 standard; peptide; 31 AA.
XX
AC   AAW03902;
XX
DT   15-APR-1997  (first entry)
XX
DE   Glucagon like peptide 1 (7-37) analogue Ala26.
XX
KW   Human; glucagon like peptide; GLP-1; analogue; stimulation;
KW   pancreas; insulin; islet cell; treatment; type II diabetes.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 20
FT                   /note= "wild type Lys substituted with Ala"
FT   Misc-difference 29
FT                   /note= "optionally absent when Arg30 and Gly31 are
FT                    absent"
FT   Misc-difference 30
FT                   /note= "optionally absent when Gly31 is absent"
FT   Misc-difference 31
FT                   /note= "optionally absent"
XX
PN   US5545618-A.
XX
PD   13-AUG-1996.
XX
PF   24-JAN-1990;   90US-0468736.
XX
PR   20-SEP-1991;   91US-0762768.
PR   24-JAN-1990;   90US-0468736.
PR   10-DEC-1993;   93US-0165516.
XX
PA   (BUCK/) BUCKLEY D I.
PA   (HABE/) HABENER J F.
PA   (MALL/) MALLORY J B.
PA   (MOJS/) MOJSOV S.
XX
PI   Buckley DI,  Habener JF,  Mallory JB,  Mojsov S;
XX
DR   WPI; 1996-383697/38.
XX
PT   New modified glucagon-like peptide I fragments - have higher
PT   activity than glucagon or have improved plasma stability, useful for
PT   treating type II diabetes
XX
PS   Example 1; page -; 16pp; English.
XX
CC   The present peptide is a specific example of a claimed human
CC   glucagon like peptide 1 (GLP-1) analogue, which is useful for
CC   stimulating insulin release from pancreatic islet cells, especially
CC   in the treatment of type II diabetes at doses of 1 pg/kg to
CC   1 mg/kg.
XX
SQ   Sequence   31 AA;

  Query Match             95.0%;  Score 134;  DB 17;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 5.7e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAAEFIAWLVKGR 30


RESULT 3
AAR45435
ID   AAR45435 standard; protein; 30 AA.
XX
AC   AAR45435;
XX
DT   27-JUN-1994  (first entry)
XX
DE   Insulinotropin derivative.
XX
KW   Insulinotropic; activity; enhancing insulin activity; treatment;
KW   Type II diabetes.
XX
OS   Synthetic.
XX
PN   WO9325579-A.
XX
PD   23-DEC-1993.
XX
PF   14-APR-1993;   93WO-US03388.
XX
PR   15-JUN-1992;   92US-0899073.
XX
PA   (PFIZ ) PFIZER INC.
XX
PI   Andrews GC,  Daumy GO,  Francoeur ML,  Larson ER;
XX
DR   WPI; 1994-007457/01.
XX
PT   New derivs. of glucagon-like peptide 1 and insulinotropin - used for
PT   enhancing insulin action in a mammal, partic. by iontophoretic admin.
XX
PS   Claim 3; Page 20; 32pp; English.
XX
CC   The sequence is that of a derivative of insulinotropin which
CC   has insulinotropic activity and is useful for enhancing insulin
CC   action in a mammal, partic. for treating Type II diabetes
CC   (claimed). It is partic. suited for delivery to a mammal by
CC   ionophoresis.
XX
SQ   Sequence   30 AA;

  Query Match             94.3%;  Score 133;  DB 15;  Length 30;
```

```
  Best Local Similarity   86.7%;  Pred. No. 8.1e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 4
AAR63247
ID   AAR63247 standard; peptide; 30 AA.
XX
AC   AAR63247;
XX
DT   02-MAY-1995  (first entry)
XX
DE   Insulinotropin (GLP-1(7-36)) for use in treating NIDDM.
XX
KW   insulinotropic activity; GLP-1; glucagon-like protein 1; NIDDM;
KW   non-insulin dependent diabetes mellitus; insulintropin; truncated.
XX
OS   Synthetic.
XX
PN   EP619322-A.
XX
PD   12-OCT-1994.
XX
PF   10-FEB-1994;   94EP-0300981.
XX
PR   07-APR-1993;   93US-0044133.
XX
PA   (PFIZ ) PFIZER INC.
PA   (PFIZ ) PFIZER CORP.
XX
PI   Danley DE,  Gelfand RA,  Geoghegan KF,  Kim Y,  Lambert WJ;
PI   Qi H,  Oih,  Hong Q,  Yesook K;
XX
DR   WPI; 1994-311774/39.
XX
PT   Treatment of non-insulin dependent diabetes mellitus - using a
PT   glucagon-like peptide 1 or deriv. with prolonged action for
PT   sustained glycaemic control
XX
PS   Claim 2; Page 46; 70pp; English.
XX
CC   This peptide is GLP-1(7-36) [GLP = glucagon-like peptide], a truncated
CC   deriv. of GLP-1. GLP-1 and its deriv.s are useful in the treatment of
CC   Non-Insulin Dependent Diabetes Mellitus (NIDDM). During processing in
CC   the pancreas and intestine, GLP-1 (AAR63245) is converted to a 31 amino
CC   acid peptide having amino acids 7-37 of GLP-1, alternatively referred
CC   to as insulinotropin. GLP-1(7-37) has insulinotropic activity, ie. it
CC   is able to stimulate, or cause to be stimulated, the synthesis of the
CC   hormone insulin. Other derivs. of GLP-1 are shown in AAR63246-51. It
CC   has been discovered that prolonged plasma elevations of GLP-1, and
CC   related polypeptides, are necessary during the meal and beyond to
CC   achieve sustained glycemic control in patients with NIDDM. The invention
CC   provides a compsn. that has prolonged action after each administration.
XX
SQ   Sequence   30 AA;

  Query Match             94.3%;  Score 133;  DB 15;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 8.1e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 5
AAR69063
ID   AAR69063 standard; peptide; 30 AA.
XX
AC   AAR69063;
XX
DT   23-AUG-1995  (first entry)
XX
DE   Amidated Glucagon like peptide 1 (GLP1) (7-36)-NH2.
XX
KW   Glucagon Like Peptide; GLP; transpeptidation; endopeptidase;
KW   trypsin; thrombin; cleavage.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   30
FT                   /label= Arg-NH2
XX
PN   WO9503405-A.
XX
PD   02-FEB-1995.
XX
PF   19-JUL-1994;   94WO-US08125.
XX
PR   20-JUL-1993;   93US-0095162.
XX
PA   (BION-) BIONEBRASKA INC.
XX
PI   Henriksen D,  Manning S,  Partridge B,  Stout J,  Wagner FW;
XX
DR   WPI; 1995-075233/10.
XX
PT   Transpeptidation of recombinant polypeptides - using
PT   endopeptidase such as trypsin or thrombin to modify C-terminal
PT   residue.
XX
PS   Claim 33; Page 50; 69pp; English.
XX
CC   The naturally occurring sequence of Glucagon Like Peptide 1 (GLP1)
CC   is AAR69072. It is a 36 AA peptide that has been recombinantly
CC   produced but without a mechanism for providing for the amidation of
CC   the C-terminal Arg residue. Amidated recombinant GLP1 (7-36)NH2
CC   (AAR69063) was prepd. from a multicopy fusion protein contg. four
CC   copies of a modified truncated GLP peptide having AA residues 7-34
CC   of the native polypeptide and the terminal AA residues A-F-A at
CC   residues 35-37 (GLP1 (7-34)-A-F-A) (AAR69064). The recombinant GLP1 (7-
CC   34)-A-F-A can be transpeptidated to yield the modified recombinant
CC   native GLP1 (7-36)-NH2 (AAR69063) as follows. Trypsin was used to
CC   cleave the peptide at the Lys-Ala bond in the presence of either
CC   Gly-Arg-NH2 or Gly-Arg-Gly addition units so that the cleavage of
CC   the Ala-Phe-Arg leaving unit is followed by the addition of
CC   Gly-Arg-NH2 or Gly-Arg-Gly to the core GLP1 (7-34) to yield either
CC    amidated 7-36 GLP1-NH2 or GLP1 7-36 with a terminal Gly (AAR69065).
XX
SQ   Sequence   30 AA;

  Query Match             94.3%;  Score 133;  DB 16;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 8.1e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 6
AAR79809
ID   AAR79809 standard; peptide; 30 AA.
XX
AC   AAR79809;
XX
DT   01-FEB-1996  (first entry)
XX
DE   Glucagon like peptide GLP-1 (7-36)amide.
XX
KW   Glucagon like peptide; GLP-1 (7-36)amide; type II diabetes;
```

KW non-insulin dependent; divalent metal cation; zinc.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 30
FT /note= "amidated"
XX
PN EP658568-A1.
XX
PD 21-JUN-1995.
XX
PF 02-DEC-1994; 94EP-0308950.
XX
PR 09-DEC-1993; 93US-0164277.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Galloway JA, Hoffmann JA;
XX
DR WPI; 1995-217011/29.
XX
PT New divalent metal complexes of glucagon-like peptide 1 - useful for
PT treating type II diabetes
XX
PS Claim 4; Page 4; 10pp; English.
XX
CC AAR79809 is the glucagon like peptide GLP-1 (7-36)amide. When
CC complexed to a divalent metal cation (pref. zinc) it can be
CC used to treat type II (non-insulin dependent) diabetes.
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 16; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
| |||||||||||| |||| |||||||||
Db 1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30

RESULT 7
AAR80548
ID AAR80548 standard; peptide; 30 AA.
XX
AC AAR80548;
XX
DT 28-FEB-1996 (first entry)
XX
DE Human glucagon like peptide (GLP-1).
XX
KW Exendin-4; diabetes mellitus; hyperglycaemia;
KW insulinotropic peptide; glucagon like peptide; GLP-1.
XX
OS Homo sapiens.
XX
PN US5424286-A.
XX
PD 13-JUN-1995.
XX
PF 24-MAY-1993; 93US-0066480.
XX
PR 24-MAY-1993; 93US-0066480.
XX
PA (ENGJ/) ENG J.
XX
PI Eng J;
XX
DR WPI; 1995-262627/34.
XX
PT Stimulating/inhibiting insulin release with exendin polypeptide(s) -
PT for treating diabetes mellitus and preventing hyperglycaemia.
XX
PS Disclosure; Columns 5-6; 17pp; English.
XX
CC AAR80548 is the human glucagon like peptide (GLP-1), to which the
CC Heloderma horridum/suspectum exendin-3/-4 peptides are analogous.
CC The exendin peptides are insulinotropic, and can therefore be used
CC in the treatment of diabetes mellitus (types I or II), and for the
CC prevention of hyperglycaemia.
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 16; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
| |||||||||||| |||| |||||||||
Db 1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30

RESULT 8
AAR98956
ID AAR98956 standard; peptide; 30 AA.
XX
AC AAR98956;
XX
DT 15-JAN-1997 (first entry)
XX
DE Target peptide (GLP1(7-36)) used in fusion protein construct.
XX
KW Fusion protein construct; isolation; purification;
KW growth hormone releasing factor; glucagon-like peptide 1;
KW parathyroid hormone; inclusion body; carbonic anhydrase.
XX
OS Synthetic.
XX
PN WO9617942-A1.
XX
PD 13-JUN-1996.
XX
PF 07-DEC-1995; 95WO-US15800.
XX
PR 07-DEC-1994; 94US-0350530.
XX
PA (BION-) BIONEBRASKA INC.
XX
PI De LA MOTTE RS, Henriksen DB, Holmquist B, Manning SD;
PI Partridge BE, Stout JS, Wagner FW;
XX
DR WPI; 1996-287186/29.
XX
PT Isolation and purificn of peptide(s) from fusion protein constructs
PT - which include a carbonic anhydrase and a variable fused
PT polypeptide
XX
PS Claim 58; Page 50; 67pp; English.
XX
CC A new method for the isolation and/or purification of a recombinant
CC peptide employs a fusion protein construct (FPC) comprising a
CC carbonic anhydrase and a variable fused polypeptide containing a
CC target peptide. The method comprises precipitating either the FPC or
CC a fragment of the FPC including the carbonic anhydrase. An
CC alternative method of producing the peptide comprises expressing the
CC FPC as part of an inclusion body. The target peptides of the FPC are
CC derived from growth hormone releasing factor (GRF), glucagon-like
CC peptide 1 (GLP1) or parathyroid hormone (PTH). This sequence
CC corresponds to amino acids 7-36 of GLP1.
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 17; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 9
AAR98975
ID   AAR98975 standard; Peptide; 30 AA.
XX
AC   AAR98975;
XX
DT   03-DEC-1996  (first entry)
XX
DE   GLP1(7-35)-NH2.
XX
KW   GLP1; C-amide; C-amidated peptide; alpha-carboxamide;
KW   recombinant protein; fusion protein; transpeptidation.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   30
FT                   /note= "C-terminal amide"
FT
XX
PN   WO9617941-A2.
XX
PD   13-JUN-1996.
XX
PF   07-DEC-1995;   95WO-US15799.
XX
PR   07-DEC-1994;   94US-0350528.
XX
PA   (BION-) BIONEBRASKA INC.
XX
PI   Heriksen DB,  Holmquist B,  Patridge BE,  Stout JS;
PI   Wagner FW;
XX
DR   WPI; 1996-287185/29.
XX
PT   Production of C-terminal alpha-carboxamidated peptide(s) - by
PT   cleavage and transpeptidation of recombinant multicopy peptide(s) or
PT   fusion constructs
XX
PS   Example 16; Page 69; 93pp; English.
XX
CC   Amidated recombinant GLP1(7-36)-NH2 (AAR98975) may be prepd. from
CC   a recombinant multicopy fusion peptide by cleavage, transamidation
CC   and photochemical rearrangement.  A DNA construct is formed by
CC   joining 4 copies of the coding sequence for GLP1(7-36)-Met
CC   (AAR98976) and a linker peptide including a thrombin cleavage site.
CC   Expression in E. coli, followed by thrombin and CNBr digestion yields
CC   GLP1(7-36)-Hse (AAR98977), which is subjected to transamidation and
CC   UV irradiation to yield GLP1(7-36)-NH2.  The amidated peptide may also
CC   be produced via GLP1(7-35)-Met (AAR98978) using a transpeptidation
CC   reaction.
XX
SQ   Sequence   30 AA;

  Query Match             94.3%;  Score 133;  DB 17;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 8.1e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 10
AAW16383
ID   AAW16383 standard; Peptide; 30 AA.
XX
AC   AAW16383;
XX
DT   01-OCT-1997  (first entry)
XX
DE   Glucagon-like peptide-1(7-36).
XX
KW   Glucagon-like peptide-1(7-36); GLP-1 (7-36); insulin secretagogue;
KW   insulinotropic hormone; type II diabetes mellitus; therapy.
XX
OS   Rattus sp.
XX
PN   US5614492-A.
XX
PD   25-MAR-1997.
XX
PF   05-MAY-1986;   86US-0859928.
XX
PR   05-SEP-1991;   91US-0756215.
PR   05-MAY-1986;   86US-0859928.
PR   26-JAN-1988;   88US-0148517.
PR   01-JUN-1990;   90US-0532111.
PR   23-NOV-1993;   93US-0156800.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Habener JF;
XX
DR   WPI; 1997-201513/18.
XX
PT   Glucagon-like peptide-1 fragment comprising amino acids 7-36 -
PT   useful for enhancing insulin production in pancreatic islet cells,
PT   especially for treating type II diabetes mellitus
XX
PS   Claim 1; Column 34; 37pp; English.
XX
CC   Glucagon-like peptide-1 (7-36) (AAW16383) comprises amino acid
CC   residues 7-36 of rat glucagon-like peptide-1 (GLP-1) (see also
CC   AAW16384).  It is naturally produced from GLP-1 in the intestine
CC   and to a lesser extent in the pancreas.  GLP-1(7-36) has
CC   insulinotropic activity, being able to stimulate the synthesis
CC   and secretion of insulin from the pancreas.  It can be produced
CC   by chemical synthesis or by proteolytic digestion of GLP-1 for use
CC   as an insulin secretagogue or for the treatment of type II diabetes
CC   mellitus.
XX
SQ   Sequence   30 AA;

  Query Match             94.3%;  Score 133;  DB 18;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 8.1e-15;
  Matches   26;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
           | ||||||||||||| ||||  |||||||||
Db       1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30


RESULT 11
AAW29555
ID   AAW29555 standard; peptide; 30 AA.
XX
AC   AAW29555;
XX
DT   02-FEB-1999  (first entry)
XX
DE   Thr-8 GLP-1(7-36)-NH2, a new glucagon-like peptide-1 analogue.
XX
KW   Glucagon-like peptide-1; GLP-1; analogue; diabetes mellitus; insulin.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   30
FT                   /note= "Arg-NH2"
XX
```

PN WO9843658-A1.
XX
PD 08-OCT-1998.
XX
PF 25-MAR-1998; 98WO-US05945.
XX
PR 31-MAR-1997; 97US-0041167.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Hoffmann JA;
XX
DR WPI; 1998-568263/48.
XX
PT New glucagon-like peptides for stimulating insulin secretion - have
PT long-lasting action without risk of inducing hypoglycaemia, useful
PT for treatment of diabetes
XX
PS Claim 8; Page 23; 26pp; English.
XX
CC The sequence represents a specifically claimed example of new
CC GLP-1 (7-37) analogues (GLP = glucagon-like peptide) which stimulate
CC insulin release and so are useful for treating diabetes mellitus.
CC The analogues have a longer-lasting action than known GLP-1 peptides
CC (they are more resistant to peptidases) and can provide overnight
CC glycaemic control without the risk of inducing hypoglycaemia.
CC (N.B. This sequence is not shown explicitly in the patent but was
CC constructed using the generic sequence of AAW29551 as a template.)
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 19; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
| ||||||||||||| |||| |||||||||
Db 1 HTEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30

RESULT 12
AAW63288
ID AAW63288 standard; peptide; 30 AA.
XX
AC AAW63288;
XX
DT 29-SEP-1998 (first entry)
XX
DE Glucagon-like peptide-1 (7-36) amide.
XX
KW GLP-1; glucagon-like peptide; obesity.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 30
FT /note= "C-terminal amide"
XX
PN WO9819698-A1.
XX
PD 14-MAY-1998.
XX
PF 04-NOV-1997; 97WO-US20114.
XX
PR 30-OCT-1997; 97US-0961405.
PR 05-NOV-1996; 96US-0030213.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI DiMarchi RD, Efendic S;
XX
DR WPI; 1998-286595/25.
XX
PT Use of glucagon-like peptide-1 and analogues and derivatives - to
PT reduce body weight, e.g., in treatment of obesity
XX
PS Claim 12; Page 18; 42pp; English.
XX
CC The patent describes a new method of reducing body weight which
CC comprises administration of a composition comprising: (i) glucagon-
CC like peptide-1 (GLP-1); (ii) a GLP-1 analogue; (iii) a GLP-1 derivative;
CC (iv) an agonist of the GLP-1 receptor; (v) an agonist of the GLP-1
CC signal transduction cascade; (vi) a compound which stimulates synthesis
CC of endogenous GLP-1; (vii) a compound that stimulates release of
CC endogenous GLP-1; or (viii) a salt of a material described in (i)-(vii).
CC The method may be used for treatment of obesity. The present sequence,
CC GLP-1 (7-36) amide, represents a preferred GLP-1 compound which can be
CC used in the method.
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 19; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
| ||||||||||||| |||| |||||||||
Db 1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30

RESULT 13
AAW63182
ID AAW63182 standard; peptide; 30 AA.
XX
AC AAW63182;
XX
DT 16-SEP-1998 (first entry)
XX
DE GLP-1(7-36).
XX
KW Glucagon-like peptide-1; GLP-1; diabetes; lipophilic; tetradecanoyl;
KW carboxynonadecanoyl; deoxycholoyl; choloyl; lithocholoyl.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 30
FT /note= "optionally the C-terminal is in amide form"
XX
PN WO9808871-A1.
XX
PD 05-MAR-1998.
XX
PF 22-AUG-1997; 97WO-DK00340.
XX
PR 20-DEC-1996; 96DK-0001470.
PR 30-AUG-1996; 96DK-0000931.
PR 08-NOV-1996; 96DK-0001259.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Knudsen LB, Nielsen PF, Sorensen PO;
XX
DR WPI; 1998-239721/21.
XX
PT Glucagon-like peptide-1 derivatives which have lipophilic
PT substituent - exhibit protracted profiles of action relative to
PT known glucagon-like peptide-1 compounds and are useful in
PT treatment of diabetes
XX
PS Claim 36; Page -; 76pp; English.
XX
CC New derivatives of glucagon-like peptide-1 (GLP-1) and its fragments
CC and their analogues are disclosed in which at least one amino acid
CC residue of the parent peptide has a lipophilic substituent attached
CC to it. The GLP-1 fragment is preferably GLP-1(A-C) where A is 1-7 and

CC C is 35-45. The lipophilic substituent is typically tetradecanoyl,
CC carboxynonadecanoyl, deoxycholoyl, choloyl or lithocholoyl, and it
CC is attached e.g. to the epsilon-amino group of a Lys residue in the
CC peptide. The present sequence represents a preferred parent GLP-1
CC fragment to which the lipophilic substituent is to be attached.
CC GLP-1 and its analogues and fragments may be used in treatment of
CC type 1 and type 2 diabetes. Prior art analogues exhibit a high
CC clearance rate from the body, which limits their usefulness. The
CC new lipophilically substituted compounds have a protracted profile
CC of action compared with known analogues, e.g. GLP-1(7-37).
CC (N.B. The present sequence is described by name in the patent
CC specification but is not explicitly shown. It is deduced from the
CC protein sequence shown in Swiss-Prot entry P01275 using information
CC given in the patent.)
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 19; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
| ||||||||||||| |||| |||||||||
Db 1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30

RESULT 14
AAW50906
ID AAW50906 standard; peptide; 30 AA.
XX
AC AAW50906;
XX
DT 17-AUG-1998 (first entry)
XX
DE Glucagon-like peptide-1 analogue SEQ ID NO:5.
XX
KW Glucagon-like peptide-1; GLP-1 (7-37); GLP-1 analogue; surgical trauma;
KW stress; hormonal response; insulin resistance; catabolic reaction;
KW human; incretin hormone.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 30
FT /note= "amidated"
XX
PN WO9808873-A1.
XX
PD 05-MAR-1998.
XX
PF 26-AUG-1997; 97WO-US15042.
XX
PR 21-AUG-1997; 97US-0024982.
PR 30-AUG-1996; 96US-0024982.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Efendic S;
XX
DR WPI; 1998-239722/21.
XX
PT Use of glucagon-like peptide-1 and analogues and their derivatives
PT - to attenuate post-surgical catabolic changes, insulin resistance
PT and hormonal responses to stress
XX
PS Claim 1; Page 13; 42pp; English.
XX
CC The present sequence represents a glucagon-like peptide-1 (GLP-1)
CC analogue, which is used in the methods of the invention. The methods
CC are: (1) for attenuating post-surgical catabolic changes and insulin
CC resistance, comprising administering glucagon-like peptide-1 (GLP-1), a
CC GLP-1 analogue, a GLP-1 derivative, or a salt of this compound; (2) for
CC attenuating post-surgical catabolic changes and hormonal responses to
CC stress, comprising administering a compound which exerts insulino-tropic
CC activity by interacting with the same receptor (or receptors) with which
CC GLP-1, GLP-1 analogues and GLP-1 derivatives interact in exerting their
CC insulino-tropic activity, and (3) for attenuating post-surgical catabolic
CC changes and hormonal responses to stress, comprising administering a
CC compound which enhances insulin sensitivity by interacting with the same
CC receptor (or receptors) with which GLP-1, GLP-1 analogues and GLP-1
CC derivatives interact to enhance insulin sensitivity. The processes are
CC useful for improving recovery after surgery by preventing the catabolic
CC reaction and insulin resistance caused by surgical trauma and
CC exacerbated by pre-operative fasting. GLP-1's short half-life, and hence
CC the need for continuous administration, are not disadvantages, as the
CC patient is usually hospitalised before surgery, and fluids are
CC continuously administered parenterally, before, during and after surgery.
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 19; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
| ||||||||||||| |||| |||||||||
Db 1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30

RESULT 15
AAY42935
ID AAY42935 standard; peptide; 30 AA.
XX
AC AAY42935;
XX
DT 20-DEC-1999 (first entry)
XX
DE Glucagon-like peptide GLP-1 (7-36).
XX
KW Glucagon-like peptide; GLP-1; antidiabetic; anti-obesity;
KW insulinotropic; appetite suppressant.
XX
OS Homo sapiens.
XX
PN WO9943707-A1.
XX
PD 02-SEP-1999.
XX
PF 25-FEB-1999; 99WO-DK00085.
XX
PR 27-FEB-1998; 98DK-0000263.
PR 27-FEB-1998; 98DK-0000268.
PR 08-APR-1998; 98DK-0000508.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Knudsen LB, Huusfeldt PO, Nielsen PF, Madsen K;
XX
DR WPI; 1999-540561/45.
XX
PT New N-modified peptide derivatives, useful for treating diabetes,
PT insulin resistance and obesity -
XX
PS Disclosure; Page 1; 62pp; English.
XX
CC New glucagon-like peptide-1 (GLP-1) derivatives are disclosed which
CC comprise residues 7-45 of GLP-1 or a fragment thereof, preferably
CC residues 7-36, 7-37 or 7-38 or their analogues, in which (a) a
CC lipophilic substituent is attached to at least one amino acid and (b)
CC the N-terminal is substituted with a group containing an optionally
CC substituted 5- or 6-membered N-heterocycle, e.g. imidazolyl. The
CC compounds stimulate secretion of insulin, suppress secretion of
CC glucagon, suppress gastric motility and/or restore glucose compliance
CC to beta-cells. They are used to treat insulin-dependent or non-insulin-
CC dependent diabetes mellitus, insulin resistance and obesity. They have

CC a longer-lasting action than GLP-1 derivatives that lack the lipophilic
CC substituent. Some of them also exist as partially structured micelle-
CC like aggregates, so have improved solubility and stability. The present
CC sequence is a specifically preferred example of a GLP-1 analogue on
CC which the derivatives are based.
XX
SQ Sequence 30 AA;

Query Match 94.3%; Score 133; DB 20; Length 30;
Best Local Similarity 86.7%; Pred. No. 8.1e-15;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 HXEGTFTSDVSSYLXGQAAXXFIAWLVKGR 30
          | ||||||||||||| ||||  |||||||||
Db      1 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR 30
```

Search completed: February 13, 2003, 11:02:17
Job time : 36 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:10:09 ; Search time 174.5 Seconds
(without alignments)
1200.205 Million cell updates/sec

Title: US-09-091-605-4
Perfect score: 93
Sequence: 1 CATGTTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	93	100.0	93	18	AAT77297	DNA encoding gluca
	2	91.4	98.3	93	18	AAT77296	DNA encoding gluca
	3	91.4	98.3	96	13	AAQ27605	DNA encoding gluca
	4	91.4	98.3	955	18	AAT75672	Human preproglucag
	5	91.4	98.3	955	18	AAT75673	Human mutated prep
	6	91.4	98.3	955	21	AAC55763	Human preproglucag
	7	91.4	98.3	955	21	AAC55765	Mutant human prepr
	8	91.4	98.3	2356	21	AAC55775	Human growth hormo
	9	91.4	98.3	3798	23	ABV25306	Human prostate exp
	10	89.8	96.6	396	22	AAF30989	Prepro-somatostati
	11	86.4	92.9	626	24	ABQ58859	Human colon cancer
	12	81.8	88.0	895	18	AAT75669	Rat preproglucagon
	13	81.8	88.0	895	21	AAC55762	Rat preproglucagon
	14	81.8	88.0	1034	11	AAQ06255	Glucagon-like pept
	15	81.8	88.0	1034	18	AAT73216	Rat prepro-glucago
	16	81.8	88.0	1034	20	AAZ20678	Preproglucagon cod
	17	80.2	86.2	387	21	AAA51462	PCpB-RR-GLIP (R26)
	18	80.2	86.2	390	21	AAA51461	PCpB-RVR-GLIP (R26
	19	80.2	86.2	390	21	AAA51464	PCpB-QAR-GLIP (R26
	20	80.2	86.2	393	21	AAA51463	PCpB-EGGR-GLIP (R2
	21	79.2	85.2	144	13	AAQ27607	GLP-I(7-37)-hbFGF
	22	78.8	84.7	387	21	AAA51460	PCpB-AM-GLIP (R26)
	23	74.8	80.4	95	22	AAF29136	Glucagon-like pept
	24	70.6	75.9	1576	13	AAQ25339	Chicken glucagon c
	25	66.4	71.4	627	20	AAX99247	Nucleotide sequenc
	26	63.2	68.0	395	21	AAA51465	PCpB-QAR-V8-GLIP f
	27	63.2	68.0	398	21	AAA51466	PCpB-LVPR-V8-GLIP
	28	61.6	66.2	315	11	AAQ04767	Sequence encoding
	29	57	61.3	110	16	AAQ91253	Glucagon like pept
	30	57	61.3	112	16	AAQ91251	Glucagon like pept
	31	57	61.3	384	16	AAQ91259	Glucagon like pept
	32	56.6	60.9	516	17	AAT07347	pAK623 fragment.
	33	55.4	59.6	279	17	AAT34870	Plasmid pBN4:GLP(7
	34	53.8	57.8	174	17	AAT34869	Plasmid pBN4:GLP(7
	35	53	57.0	1134	22	AAC86599	DNA encoding a Bac
	36	53	57.0	1158	22	AAC86600	DNA encoding a Bac
	37	49.8	53.5	78	21	AAZ57465	Kunitz protease in
	38	49.8	53.5	78	22	AAS14247	Oligonucleotide 62
c	39	47	50.5	108	16	AAQ91252	Glucagon like pept
c	40	45.8	49.2	112	16	AAQ91254	Glucagon like pept
	41	37.8	40.6	65	24	ABN56946	Mouse spliced tran
	42	37.4	40.2	57	18	AAV02282	Oligonucleotide GL
c	43	35	37.6	70	22	AAC86614	PCR primer used to
c	44	34.2	36.8	48	18	AAV02283	Oligonucleotide GL
	45	34	36.6	492	19	AAV33163	Heloderma suspectu

ALIGNMENTS

RESULT 1
AAT77297
ID AAT77297 standard; DNA; 93 BP.
XX
AC AAT77297;
XX
DT 14-JAN-1998 (first entry)
XX
DE DNA encoding glucagon-like peptide GLP-1(7-37) Val8 analogue.
XX
KW Diabetes; non-insulin dependent diabetes mellitus; NIDDM;
KW insulin dependent diabetes mellitus; IDDM; gene therapy;
KW glucagon-like peptide; GLP; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT mat_peptide 1..93
FT /*tag= a
FT /product= Val8_GLP-1(7-37)
XX
PN WO9729180-A1.
XX
PD 14-AUG-1997.
XX
PF 06-FEB-1997; 97WO-US01978.
XX
PR 23-FEB-1996; 96GB-0003847.
PR 06-FEB-1996; 96US-0012111.
XX
PA (ELIL) LILLY & CO ELI.
XX

```
PI   Borts TL,  Broderick CL,  Dimarchi RD,  Grinnell BW;
PI   Miller AR;
XX
DR   WPI; 1997-415336/38.
DR   P-PSDB; AAW24390.
XX
PT   Gene therapy of type I and type II diabetes - by in vivo expression
PT   of glucagon like peptide GLP-1(7-37) analogue
XX
PS   Claim 8; Page 24; 31pp; English.
XX
CC   This DNA sequence encodes an analogue of a glucagon-like peptide
CC   GLP-1(7-37). This peptide provides a means of delivering long term
CC   amounts of a GLP-1(7-37)-based protein, which is useful in treating type
CC   I and type II diabetes. A stable mammalian cell line, which  is
CC   immunologically isolated from the mammal's immune system, is transformed
CC   with a vector expressing a protein of the above sequence. This
CC   transformed cell line is then implanted into the individual needing
CC   treatment. Once implanted, the GLP-1(7-37) analogue, in conjunction with
CC   high serum glucose levels, causes pancreatic cells to produce insulin in
CC   non-insulin dependent diabetes mellitus (NIDDM) and delays gastric
CC   emptying in both NIDDM and insulin dependent diabetes mellitus (IDDM)
CC   patients. An expression vector coding for a protein of the above
CC   sequence can also be directly injected into the mammal, such that the
CC   vector is incorporated into a cell and secretes the protein. This method
CC   overcomes the problems of the short serum half life of GLP-1(7-37),
CC   allowing delivery of effective long term amounts for diabetes treatment.
XX
SQ   Sequence 93 BP; 24 A; 16 C; 28 G; 25 T; 0 other;

  Query Match             100.0%;  Score 93;  DB 18;  Length 93;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-23;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 2
AAT77296
ID   AAT77296 standard; DNA; 93 BP.
XX
AC   AAT77296;
XX
DT   14-JAN-1998  (first entry)
XX
DE   DNA encoding glucagon-like peptide GLP-1(7-37) Ala8 analogue.
XX
KW   Diabetes; non-insulin dependent diabetes mellitus; NIDDM;
KW   insulin dependent diabetes mellitus; IDDM; gene therapy;
KW   glucagon-like peptide; GLP; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   mat_peptide     1..93
FT                   /*tag=  a
FT                   /product= Ala8_GLP-1(7-37)
XX
PN   WO9729180-A1.
XX
PD   14-AUG-1997.
XX
PF   06-FEB-1997;   97WO-US01978.
XX
PR   23-FEB-1996;   96GB-0003847.
PR   06-FEB-1996;   96US-0012111.
XX
```

```
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Borts TL,  Broderick CL,  Dimarchi RD,  Grinnell BW;
PI   Miller AR;
XX
DR   WPI; 1997-415336/38.
DR   P-PSDB; AAW24389.
XX
PT   Gene therapy of type I and type II diabetes - by in vivo expression
PT   of glucagon like peptide GLP-1(7-37) analogue
XX
PS   Claim 8; Page 24; 31pp; English.
XX
CC   This DNA sequence encodes an analogue of a glucagon-like peptide
CC   GLP-1(7-37). This peptide provides a means of delivering long term
CC   amounts of a GLP-1(7-37)-based protein, which is useful in treating type
CC   I and type II diabetes. A stable mammalian cell line, which  is
CC   immunologically isolated from the mammal's immune system, is transformed
CC   with a vector expressing a protein of the above sequence. This
CC   transformed cell line is then implanted into the individual needing
CC   treatment. Once implanted, the GLP-1(7-37) analogue, in conjunction with
CC   high serum glucose levels, causes pancreatic cells to produce insulin in
CC   non-insulin dependent diabetes mellitus (NIDDM) and delays gastric
CC   emptying in both NIDDM and insulin dependent diabetes mellitus (IDDM)
CC   patients. An expression vector coding for a protein of the above
CC   sequence can also be directly injected into the mammal, such that the
CC   vector is incorporated into a cell and secretes the protein. This method
CC   overcomes the problems of the short serum half life of GLP-1(7-37),
CC   allowing delivery of effective long term amounts for diabetes treatment.
XX
SQ   Sequence 93 BP; 24 A; 17 C; 28 G; 24 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 18;  Length 93;
  Best Local Similarity   98.9%;  Pred. No. 5.4e-23;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 3
AAQ27605
ID   AAQ27605 standard; DNA; 96 BP.
XX
AC   AAQ27605;
XX
DT   04-FEB-1993  (first entry)
XX
DE   DNA encoding glucagon-like peptide I (7-37).
XX
KW   Human parathyroid hormone production; osteoporosis; GLP-I (7-37);
KW   hypoparathyroidism; hypertension; insulinotropin; ss.
XX
OS   Synthetic.
XX
PN   EP499990-A.
XX
PD   26-AUG-1992.
XX
PF   15-FEB-1992;   92EP-0102543.
XX
PR   19-FEB-1991;   91JP-0024841.
PR   18-OCT-1991;   91JP-0271438.
PR   24-OCT-1991;   91JP-0277724.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Fukuda T,  Koyama N,  Kuriyama M,  Nishimura O;
```

```
XX
DR   WPI; 1992-286114/35.
XX
PT   Cysteine-free peptide prodn., e.g. human parathyroid hormone
PT   deriv. - by culturing transformant to produce a fusion protein
PT   comprising the cysteine-free peptide fused to a cysteine at its
PT   N-terminus where cleavage can occur
XX
PS   Disclosure; Page 5; 60pp; English.
XX
CC   The DNA codes for glucagon-like peptide I (7-37) [GLP-I (7-37)]
CC   (insulinotropin). PTH (1-34), it may be used in a method of culturing
CC   a transformant to produce a fusion protein comprising a cysteine-free
CC   peptide fused to a cysteine at its N-terminus where cleavage can occur.
CC   This method can be used to produce peptides which can be used as a
CC   pharmaceutical or industry in general. See also AAQ27603-Q27622.
XX
SQ   Sequence 96 BP; 24 A; 18 C; 29 G; 25 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 13;  Length 96;
  Best Local Similarity   98.9%;  Pred. No. 5.5e-23;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 4
AAT75672
ID   AAT75672 standard; DNA; 955 BP.
XX
AC   AAT75672;
XX
DT   03-FEB-1998  (first entry)
XX
DE   Human preproglucagon cDNA.
XX
KW   Recombinant protein; expression; secretory cell line; human;
KW   glucagon; peptide hormone; amidation; insulinoma; RIN; rat;
KW   diabetes; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             27..569
FT                   /*tag=  a
XX
PN   WO9726321-A2.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997;   97WO-US00761.
XX
PR   15-OCT-1996;   96US-0028427.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (BETA-) BETAGENE INC.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Clark SA,  Halban PA,  Kruse F,  McGarry D,  Newgard CB;
PI   Normington KD, Quaade C,  Thigpen AE;
XX
DR   WPI; 1997-385326/35.
DR   P-PSDB; AAW22080.
XX
PT   Recombinant cell engineered to provide amylin to a mammal - useful
PT   to treat e.g. angiogenesis, anorexia, obesity, hypertension,
PT   osteoporosis etc.
XX
PS   Example 10; Page 283-284; 336pp; English.
XX
CC   This cDNA sequence includes a coding sequence for human
CC   preproglucagon (see AAW22080).  It was produced from pancreatic
CC   cDNA by PCR amplification (see AAT75670-71), and has been ligated
CC   into pNoTA/T7, generating pNoTAT7/h.glucagon.  A mutated human
CC   preproglucagon cDNA (see AAT75673) has also been produced.  Rat
CC   insulinoma RIN cell lines expressing preproglucagon demonstrated
CC   efficient amidation of a secreted, processed polypeptide.  The
CC   invention provides methods for production of heterologous
CC   polypeptides using recombinantly engineered cell lines.  Also
CC   described are methods of engineering cells for high level
CC   expression, methods of large-scale heterologous protein production,
CC   and methods for treatment of disease in vivo using viral delivery
CC   systems and recombinant cell lines.
XX
SQ   Sequence 955 BP; 301 A; 181 C; 203 G; 270 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 18;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-22;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 5
AAT75673
ID   AAT75673 standard; DNA; 955 BP.
XX
AC   AAT75673;
XX
DT   03-FEB-1998  (first entry)
XX
DE   Human mutated preproglucagon cDNA.
XX
KW   Recombinant protein; expression; secretory cell line; human;
KW   glucagon; peptide hormone; amidation; insulinoma; RIN; rat;
KW   diabetes; gene therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             27..569
FT                   /*tag=  a
FT   mutation        181
FT                   /*tag=  b
FT                   /note= "C at position 181 creates an Ala-52 codon"
FT   mutation        204
FT                   /*tag=  b
FT                   /note= "T at position 204 creates an SpeI site"
XX
PN   WO9726321-A2.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997;   97WO-US00761.
XX
PR   15-OCT-1996;   96US-0028427.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (BETA-) BETAGENE INC.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Clark SA,  Halban PA,  Kruse F,  McGarry D,  Newgard CB;
PI   Normington KD, Quaade C,  Thigpen AE;
```

```
XX
DR   WPI; 1997-385326/35.
DR   P-PSDB; AAW22081.
XX
PT   Recombinant cell engineered to provide amylin to a mammal - useful
PT   to treat e.g. angiogenesis, anorexia, obesity, hypertension,
PT   osteoporosis etc.
XX
PS   Example 10; Page 285-286; 336pp; English.
XX
CC   This cDNA sequence includes a coding sequence for a mutated human
CC   preproglucagon (see AAW22081).  It was produced from plasmid
CC   pNoTAT7/h.glucagon (see AAT75672) using mutagenic primers.  2
CC   Mutations were created, one altering the native Arg-52 codon to an
CC   Ala codon, and one creating an SpeI site, with no change in the
CC   deduced amino acid sequence.  The mutated sequence has been used
CC   to construct plasmid pNiTAT7/mut.glucagon.  Rat insulinoma RIN cell
CC   lines expressing preproglucagon demonstrated efficient amidation of
CC   a secreted, processed polypeptide.  The invention provides methods
CC   for production of heterologous polypeptides using recombinantly
CC   engineered cell lines.  Also described are methods of engineering
CC   cells for high level expression, methods of large-scale
CC   heterologous protein production, and methods for treatment of
CC   disease in vivo using viral delivery systems and recombinant cell
CC   lines.
XX
SQ   Sequence 955 BP; 302 A; 180 C; 202 G; 271 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 18;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-22;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 6
AAC55763
ID   AAC55763 standard; cDNA; 955 BP.
XX
AC   AAC55763;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human preproglucagon encoding cDNA.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (BETA-) BETAGENE INC.
XX
PI   Newgard CB,  Halban P,  Normington KD,  Thigpen AE,  Quaade C;
PI   Kruse F,  McGarry D,  Clark SA;
XX
DR   WPI; 2000-586352/55.
DR   P-PSDB; AAB26774.
XX
PT   Producing mammalian amylin, useful for regulating blood glucose and
PT   insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT   employing recombinantly engineered secretory cell lines -
XX
PS   Example 10; Column 175-178; 136pp; English.
XX
CC   This invention relates to a method for producing mammalian amylin. The
CC   method relies on the use of a recombinantly engineered secretory cell
CC   line. The method comprises:
CC   (a) providing a starting secretory cell that has a regulated secretory
CC   pathway;
CC   (b) introducing, into the starting secretory cell, an amylin-encoding
CC   gene operatively linked to a first promoter;
CC   (c) selecting a secretory cell of (b) that exhibits increased production
CC   of biologically active amylin as compared to the starting secretory
CC   cell; and (d) culturing the selected secretory cell.
CC   Amylin is an insulin modulator, and the method results in antidiabetic,
CC   hypotensive and osteopathic activity. The amylins produced are useful
CC   for regulating blood glucose levels, as well as in modulating the
CC   circulating levels of insulin in a mammal. The amylins produced maybe
CC   used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC   disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC   requiring amylin regulation. The invention includes cDNA and protein
CC   sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC   AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC   the invention for the production of mammalian amylin.
XX
SQ   Sequence 955 BP; 301 A; 181 C; 203 G; 270 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 21;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-22;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 7
AAC55765
ID   AAC55765 standard; cDNA; 955 BP.
XX
AC   AAC55765;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Mutant human preproglucagon cDNA.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (BETA-) BETAGENE INC.
XX
PI   Newgard CB,  Halban P,  Normington KD,  Thigpen AE,  Quaade C;
```

```
PI   Kruse F,  McGarry D,  Clark SA;
XX
DR   WPI; 2000-586352/55.
DR   P-PSDB; AAB26775.
XX
PT   Producing mammalian amylin, useful for regulating blood glucose and
PT   insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT   employing recombinantly engineered secretory cell lines -
XX
PS   Example 10; Column 179-180; 136pp; English.
XX
CC   This invention relates to a method for producing mammalian amylin. The
CC   method relies on the use of a recombinantly engineered secretory cell
CC   line. The method comprises:
CC   (a) providing a starting secretory cell that has a regulated secretory
CC   pathway;
CC   (b) introducing, into the starting secretory cell, an amylin-encoding
CC   gene operatively linked to a first promoter;
CC   (c) selecting a secretory cell of (b) that exhibits increased production
CC   of biologically active amylin as compared to the starting secretory
CC   cell; and (d) culturing the selected secretory cell.
CC   Amylin is an insulin modulator, and the method results in antidiabetic,
CC   hypotensive and osteopathic activity. The amylins produced are useful
CC   for regulating blood glucose levels, as well as in modulating the
CC   circulating levels of insulin in a mammal. The amylins produced maybe
CC   used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC   disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC   requiring amylin regulation. The invention includes cDNA and protein
CC   sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC   AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC   the invention for the production of mammalian amylin.
XX
SQ   Sequence 955 BP; 302 A; 180 C; 202 G; 271 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 21;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-22;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 8
AAC55775
ID   AAC55775 standard; DNA; 2356 BP.
XX
AC   AAC55775;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human growth hormone and mutated proglucagon fusion DNA sequence.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (BETA-) BETAGENE INC.
XX
PI   Newgard CB,  Halban P,  Normington KD,  Thigpen AE,  Quaade C;
PI   Kruse F,  McGarry D,  Clark SA;
XX
DR   WPI; 2000-586352/55.
DR   P-PSDB; AAB26777.
XX
PT   Producing mammalian amylin, useful for regulating blood glucose and
PT   insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT   employing recombinantly engineered secretory cell lines -
XX
PS   Example 12; Column 187-190; 136pp; English.
XX
CC   This invention relates to a method for producing mammalian amylin. The
CC   method relies on the use of a recombinantly engineered secretory cell
CC   line. The method comprises:
CC   (a) providing a starting secretory cell that has a regulated secretory
CC   pathway;
CC   (b) introducing, into the starting secretory cell, an amylin-encoding
CC   gene operatively linked to a first promoter;
CC   (c) selecting a secretory cell of (b) that exhibits increased production
CC   of biologically active amylin as compared to the starting secretory
CC   cell; and (d) culturing the selected secretory cell.
CC   Amylin is an insulin modulator, and the method results in antidiabetic,
CC   hypotensive and osteopathic activity. The amylins produced are useful
CC   for regulating blood glucose levels, as well as in modulating the
CC   circulating levels of insulin in a mammal. The amylins produced maybe
CC   used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC   disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC   requiring amylin regulation. The invention includes cDNA and protein
CC   sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC   AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC   the invention for the production of mammalian amylin.
XX
SQ   Sequence 2356 BP; 614 A; 600 C; 581 G; 561 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 21;  Length 2356;
  Best Local Similarity   98.9%;  Pred. No. 1.4e-22;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1704 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 1763

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    1764 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 1796


RESULT 9
ABV25306
ID   ABV25306 standard; cDNA; 3798 BP.
XX
AC   ABV25306;
XX
DT   16-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 25297.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US05171.
XX
PR   17-FEB-2000; 2000US-183319P.
PR   16-MAR-2000; 2000US-189862P.
```

```
PR   25-MAY-2000; 2000US-207454P.
PR   09-JUN-2000; 2000US-211314P.
PR   18-JUL-2000; 2000US-219007P.
PR   13-DEC-2000; 2000US-255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer  -
XX
PS   Claim 1; Page 4978-4979; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
CC   (g) determining whether prostate cancer has metastasized in a patient;
CC   (h) assessing the aggressiveness or indolence of prostate cancer in a
CC   patient;
CC   (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ   Sequence 3798 BP; 1166 A; 563 C; 616 G; 1364 T; 89 other;

  Query Match             98.3%;  Score 91.4;  DB 23;  Length 3798;
  Best Local Similarity   98.9%;  Pred. No. 1.6e-22;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3411 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 3470

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    3471 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 3503


RESULT 10
AAF30989
ID   AAF30989 standard; DNA; 396 BP.
XX
AC   AAF30989;
XX
DT   23-JUL-2001  (first entry)
XX
DE   Prepro-somatostatin/GLP-1 coding sequence of pXIT-39.
XX
KW   Somatostatin; glucagon-like peptide 1; GLP-1; antidiabetic;
KW   drug delivery; diabetes; gene therapy; pXIT-38; ss.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..390
FT                   /*tag= a
FT                   /product= "prepro-somatostatin/cleavage site/GLP-1
FT                            fusion"
XX
PN   WO200136643-A1.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US31634.
XX
PR   19-NOV-1999;   99US-0166508.
XX
PA   (TRAN-) TRANSKARYOTIC THERAPIES INC.
XX
PI   Treco DA,  Concino MF,  Duguay SJ;
XX
DR   WPI; 2001-355636/37.
XX
PT   New nucleic acid constructs useful for transforming cells useful as a
PT   drug delivery vehicle -
XX
PS   Example 1; Fig 3; 89pp; English.
XX
CC   The present sequence is that of a DNA sequence encoding a
CC   prepro-somatostatin-glucagon-like peptide 1 (GLP-1) fusion protein.
CC   The coding region is flanked by a 5' BamHI site and a 3' XhoI site.
CC   A multibasic cleavage site separates the prepro-somatostatin and
CC   GLP-1 moieties of the fusion protein.  The DNA sequence in plasmid
CC   pXIT-39 was designed for the stable, constitutive expression of
CC   GLP-1 (7-37) in human cells.  This is an example of nucleic acid
CC   constructs of the invention designed for the expression of small
CC   peptides.  The small peptides are especially therapeutic peptides
CC   such as small hormones and antidiabetic peptides such as GLP-1,
CC   exedin-4 and gastric inhibitory polypeptide.  Claimed methods
CC   of treating a subject having diabetes involve administering the
CC   nucleic acid construct or a cell capable of expressing the small
CC   peptide.  Transfected primary or secondary cells or cell strains
CC   have wide applicability as vehicles or delivery systems for
CC   therapeutic proteins.  By controlling the number of cells introduced
CC   into an individual, one can control the amount of the protein
CC   delivered to the body.  In addition, in some cases, it is possible
CC   to remove the transfected cells if there is no longer a need for
CC   the product.  Human fibroblasts transfected with pXIT-39 secreted
CC   human GLP-1(7-37) and GLP-1(3-36) (see AAB82335-36).
XX
SQ   Sequence 396 BP; 80 A; 112 C; 121 G; 83 T; 0 other;

  Query Match             96.6%;  Score 89.8;  DB 22;  Length 396;
  Best Local Similarity   97.8%;  Pred. No. 3.1e-22;
  Matches   91;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     286 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 345

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||| ||
Db     346 GAATTCATTGCTTGGCTGGTGAAAGGCCGACGA 378


RESULT 11
ABQ58859
ID   ABQ58859 standard; cDNA; 626 BP.
XX
AC   ABQ58859;
XX
DT   02-AUG-2002  (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:2554.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200229086-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30732.
```

```
XX
PR   02-OCT-2000; 2000US-237271P.
XX
PA   (FARB ) BAYER CORP.
XX
PI   Burgess C,  Astle JH,  Carroll E,  Catino TJ,  Dwivedi P,  Molino GA;
PI   Thiaglingam A,  Lewis ME;
XX
DR   WPI; 2002-426115/45.
XX
PT   New isolated nucleic acid that is differentially expressed in cancer
PT   tissues useful for determining the presence of colon cancer in a cell
PT   or tissue type, and in antisense therapy  -
XX
PS   Claim 1; Fig 1; 796pp; English.
XX
CC   ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC   expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC   encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC   used in antisense therapy. An antibody immunoreactive with a polypeptide
CC   encoded by (I) is useful for detecting cancer in a patient sample, and
CC   for detecting the presence or absence of a polynucleotide encoded by a
CC   nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC   from (I) can be used for determining the presence of a nucleic acid which
CC   hybridises to (I), and for determining the phenotype of cells in a sample
CC   of cells from a patient. (I) is useful for determining the presence of
CC   colon cancer in a cell or tissue type, for determining the presence or
CC   state of other type of cancer, in antisense therapy, to generate
CC   macroarrays on a solid surface, to identify a chromosome on which the
CC   corresponding gene resides, and in tissue profiling, forensics, genetic
CC   analysis, mapping and diagnostic applications. (I) can be used to raise
CC   antibodies, and to screen for peptide analogues and antagonists.
XX
SQ   Sequence 626 BP; 170 A; 134 C; 143 G; 159 T; 20 other;

  Query Match             92.9%;  Score 86.4;  DB 24;  Length 626;
  Best Local Similarity   98.9%;  Pred. No. 5.8e-21;
  Matches   87;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 312

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCC 88
           ||||||||||||||||||||||||||||
Db     313 GAATTCATTGCTTGGCTGGTGAAAGGCC 340


RESULT 12
AAT75669
ID   AAT75669 standard; DNA; 895 BP.
XX
AC   AAT75669;
XX
DT   03-FEB-1998  (first entry)
XX
DE   Rat preproglucagon cDNA.
XX
KW   Recombinant protein; expression; secretory cell line; rat;
KW   glucagon; peptide hormone; amidation; insulinoma; RIN; diabetes;
KW   gene therapy; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             52..594
FT                   /*tag=  a
XX
PN   WO9726321-A2.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997;   97WO-US00761.
XX
PR   15-OCT-1996;   96US-0028427.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (BETA-) BETAGENE INC.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Clark SA,  Halban PA,  Kruse F,  McGarry D,  Newgard CB;
PI   Normington KD, Quaade C,  Thigpen AE;
XX
DR   WPI; 1997-385326/35.
DR   P-PSDB; AAW22079.
XX
PT   Recombinant cell engineered to provide amylin to a mammal - useful
PT   to treat e.g. angiogenesis, anorexia, obesity, hypertension,
PT   osteoporosis etc.
XX
PS   Example 10; Page 281-282; 336pp; English.
XX
CC   This cDNA sequence includes a coding sequence for rat
CC   preproglucagon (see AAW22079).  It was produced from pancreatic
CC   cDNA by PCR amplification (see AAT75667-68), and has been ligated
CC   into pNoTA/T7, generating pNoTAT7/r.Glucagon.  Rat insulinoma RIN
CC   cell lines expressing preproglucagon demonstrated efficient
CC   amidation of a secreted, processed polypeptide.  The invention
CC   provides methods for production of heterologous polypeptides using
CC   recombinantly engineered cell lines.  Also described are methods of
CC   engineering cells for high level expression, methods of large-scale
CC   heterologous protein production, and methods for treatment of
CC   disease in vivo using viral delivery systems and recombinant cell
CC   lines.
XX
SQ   Sequence 895 BP; 266 A; 211 C; 200 G; 218 T; 0 other;

  Query Match             88.0%;  Score 81.8;  DB 18;  Length 895;
  Best Local Similarity   92.5%;  Pred. No. 2.8e-19;
  Matches   86;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     343 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 402

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     403 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 435


RESULT 13
AAC55762
ID   AAC55762 standard; cDNA; 895 BP.
XX
AC   AAC55762;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Rat preproglucagon encoding cDNA.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Rattus sp.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
```

```
PA   (BETA-) BETAGENE INC.
XX
PI   Newgard CB,  Halban P,  Normington KD,  Thigpen AE,  Quaade C;
PI   Kruse F,  McGarry D,  Clark SA;
XX
DR   WPI; 2000-586352/55.
DR   P-PSDB; AAB26773.
XX
PT   Producing mammalian amylin, useful for regulating blood glucose and
PT   insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT   employing recombinantly engineered secretory cell lines -
XX
PS   Example 10; Column 173-174; 136pp; English.
XX
CC   This invention relates to a method for producing mammalian amylin. The
CC   method relies on the use of a recombinantly engineered secretory cell
CC   line. The method comprises:
CC   (a) providing a starting secretory cell that has a regulated secretory
CC   pathway;
CC   (b) introducing, into the starting secretory cell, an amylin-encoding
CC   gene operatively linked to a first promoter;
CC   (c) selecting a secretory cell of (b) that exhibits increased production
CC   of biologically active amylin as compared to the starting secretory
CC   cell; and (d) culturing the selected secretory cell.
CC   Amylin is an insulin modulator, and the method results in antidiabetic,
CC   hypotensive and osteopathic activity. The amylins produced are useful
CC   for regulating blood glucose levels, as well as in modulating the
CC   circulating levels of insulin in a mammal. The amylins produced maybe
CC   used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC   disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC   requiring amylin regulation. The invention includes cDNA and protein
CC   sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC   AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC   the invention for the production of mammalian amylin.
XX
SQ   Sequence 895 BP; 266 A; 212 C; 199 G; 218 T; 0 other;

  Query Match             88.0%;  Score 81.8;  DB 21;  Length 895;
  Best Local Similarity   92.5%;  Pred. No. 2.8e-19;
  Matches   86;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     343 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 402

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     403 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 435


RESULT 14
AAQ06255
ID   AAQ06255 standard; DNA; 1034 BP.
XX
AC   AAQ06255;
XX
DT   29-JAN-1991  (first entry)
XX
DE   Glucagon-like peptide, GLP-1 (7-37).
XX
KW   Insulin; diabetes mellitus; insulinotropic: pancreatic beta cells.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             61..603
FT                   /*tag=  a
FT   sig_peptide     61..120
FT                   /*tag=  b
XX
PN   WO9011296-A.
XX
PD   04-OCT-1990.
XX
PF   20-MAR-1989;   89WO-US01121.
XX
PR   20-MAR-1989;   89WO-US01121.
XX
PA   (GEHO-) GEN HOSPITAL CORP.
XX
PI   Habener JF;
XX
DR   WPI; 1990-320226/42.
XX
PT   New glucagon-like peptide (GLP-1) - having insulin
PT   formation-stimulating activity and useful in treating diabetes
PT   mellitus.
XX
PS   Claim 6; Page 39; 52pp; English.
XX
CC   The peptide has insulinotropic activity specifically for pancreatic
CC   beta cells. The peptide is derived from glucagon which, after
CC   synthesis is cleaved into three peptides: glucgon, glucagon-like
CC   peptide 1 (GLP-1) and GLP-2.  GLP-1 has 37 AAs in its unprocessed
CC   form and is unable to mediate the induction of insulin biosynthesis.
CC   It is, however, naturally converted to a 31 AA-long peptide having
CC   AAs 7-37 of GLP-1.  Preferred derivs. have an H2 gp at the
CC   N-terminal and an OH, OM, or NR'R'' gp at the C-terminal where M= a
CC   cation or lower alkyl gp., and R' and R''= H or a lower alkyl gp.
CC   Prepns. contg. the peptide or derivs. are useful in the study of
CC   the pathogenesis of maturity onset of diabetes mellitus and also in
CC   therapy.
CC   See also AAR07398.
XX
SQ   Sequence 1034 BP; 315 A; 238 C; 226 G; 255 T; 0 other;

  Query Match             88.0%;  Score 81.8;  DB 11;  Length 1034;
  Best Local Similarity   92.5%;  Pred. No. 2.9e-19;
  Matches   86;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     352 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 411

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     412 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 444


RESULT 15
AAT73216
ID   AAT73216 standard; DNA; 1034 BP.
XX
AC   AAT73216;
XX
DT   01-OCT-1997  (first entry)
XX
DE   Rat prepro-glucagon DNA.
XX
KW   Glucagon-like peptide-1(7-36); GLP-1; insulin secretagogue;
KW   insulinotropic hormone; type II diabetes mellitus; therapy; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             61..603
FT                   /*tag=  a
FT   sig_peptide     61..120
FT                   /*tag=  b
FT   mat_peptide     121..600
FT                   /*tag=  c
XX
PN   US5614492-A.
XX
PD   25-MAR-1997.
XX
```

PF 05-MAY-1986; 86US-0859928.
XX
PR 05-SEP-1991; 91US-0756215.
PR 05-MAY-1986; 86US-0859928.
PR 26-JAN-1988; 88US-0148517.
PR 01-JUN-1990; 90US-0532111.
PR 23-NOV-1993; 93US-0156800.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Habener JF;
XX
DR WPI; 1997-201513/18.
DR P-PSDB; AAW16384.
XX
PT Glucagon-like peptide-1 fragment comprising amino acids 7-36 -
PT useful for enhancing insulin production in pancreatic islet cells,
PT especially for treating type II diabetes mellitus
XX
PS Disclosure; Fig 1-1A; 37pp; English.
XX
CC A DNA molecule (AAT73216) codes for rat preproglucagon (AAW16384), a
CC polypeptide that is processed by proteolytic cleavage to glucagon,
CC glucagon-like peptide-1 (GLP-1) and GLP-2. A peptide fragment
CC of GLP-1, GLP-1(7-36) (see also AAW16363), has been shown to possess
CC hormonal activity and can be used as an insulin secretagogue and
CC for the treatment of type II diabetes mellitus.
XX
SQ Sequence 1034 BP; 316 A; 237 C; 226 G; 255 T; 0 other;

Query Match 88.0%; Score 81.8; DB 18; Length 1034;
Best Local Similarity 92.5%; Pred. No. 2.9e-19;
Matches 86; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     352 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 411

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     412 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 444
```

Search completed: February 14, 2003, 08:03:04
Job time : 175.5 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:55:04 ; Search time 48 Seconds
(without alignments)
594.186 Million cell updates/sec

Title: US-09-091-605-4
Perfect score: 93
Sequence: 1 CATGTTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	91.4	98.3	528	2	US-08-835-231-8	Sequence 8, Appli
	2	91.4	98.3	528	4	US-09-108-661-8	Sequence 8, Appli
	3	91.4	98.3	955	3	US-08-784-582-57	Sequence 57, Appl
	4	91.4	98.3	955	3	US-08-784-582-60	Sequence 60, Appl
	5	91.4	98.3	2356	3	US-08-784-582-72	Sequence 72, Appl
	6	89.8	96.6	528	2	US-08-835-231-7	Sequence 7, Appli
	7	89.8	96.6	528	4	US-09-108-661-7	Sequence 7, Appli
	8	81.8	88.0	895	3	US-08-784-582-55	Sequence 55, Appl
	9	79.2	85.2	144	2	US-08-835-231-17	Sequence 17, Appl
	10	79.2	85.2	144	4	US-09-108-661-17	Sequence 17, Appl
	11	49.8	53.5	78	2	US-08-829-876-22	Sequence 22, Appl
	12	49.8	53.5	78	4	US-09-234-874A-22	Sequence 22, Appl
	13	37.4	40.2	57	2	US-08-811-028-43	Sequence 43, Appl
c	14	34.2	36.8	48	2	US-08-811-028-44	Sequence 44, Appl
c	15	32.8	35.3	53	2	US-08-811-028-46	Sequence 46, Appl
c	16	30.6	32.9	73	2	US-08-829-876-24	Sequence 24, Appl
c	17	30.6	32.9	73	4	US-09-234-874A-24	Sequence 24, Appl
c	18	29.4	31.6	52	2	US-08-829-876-23	Sequence 23, Appl
c	19	29.4	31.6	52	4	US-09-234-874A-23	Sequence 23, Appl
	20	28	30.1	2409	1	US-08-392-828C-1	Sequence 1, Appli
	21	28	30.1	2409	3	US-09-330-945-1	Sequence 1, Appli
	22	27.8	29.9	5558	4	US-08-961-527-103	Sequence 103, App
	23	26.2	28.2	36	2	US-08-811-028-45	Sequence 45, Appl
c	24	25.8	27.7	144	2	US-08-835-231-17	Sequence 17, Appl
c	25	25.8	27.7	144	4	US-09-108-661-17	Sequence 17, Appl
	26	25.4	27.3	3416	2	US-08-701-240-3	Sequence 3, Appli
	27	25.4	27.3	3416	3	US-09-138-236-3	Sequence 3, Appli
c	28	25	26.9	725	4	US-09-221-017B-198	Sequence 198, App
	29	24.8	26.7	169998	4	US-09-676-610B-24	Sequence 24, Appl
	30	24.2	26.0	516	1	US-08-532-828B-13	Sequence 13, Appl
	31	24.2	26.0	1263	1	US-08-532-828B-11	Sequence 11, Appl
	32	24.2	26.0	1643	1	US-08-532-828B-1	Sequence 1, Appli
	33	24.2	26.0	1643	1	US-08-532-828B-7	Sequence 7, Appli
	34	24.2	26.0	1643	1	US-08-532-828B-9	Sequence 9, Appli
	35	24.2	26.0	1643	1	US-08-700-359-7	Sequence 7, Appli
	36	24.2	26.0	1643	1	US-08-700-359-8	Sequence 8, Appli
	37	24.2	26.0	1643	1	US-08-700-359-10	Sequence 10, Appl
	38	24.2	26.0	1643	1	US-08-674-168-21	Sequence 21, Appl
	39	24.2	26.0	1643	2	US-08-596-366-5	Sequence 5, Appli
	40	24.2	26.0	1643	2	US-08-596-366-7	Sequence 7, Appli
	41	24.2	26.0	1643	2	US-08-967-104-5	Sequence 5, Appli
	42	24.2	26.0	1643	2	US-08-967-104-7	Sequence 7, Appli
	43	24.2	26.0	1643	3	US-08-985-908-3	Sequence 3, Appli
	44	24.2	26.0	1643	3	US-08-985-908-4	Sequence 4, Appli
	45	24.2	26.0	1643	3	US-08-985-908-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-835-231-8
; Sequence 8, Application US/08835231
; Patent No. 5861284
; GENERAL INFORMATION:
; APPLICANT: NISHIMURA, Osamu
; APPLICANT: KURIYAMA, Masato
; APPLICANT: KOYAMA, No. 5861284uyuki
; APPLICANT: FUKUDA, Tsunehiko
; TITLE OF INVENTION: METHOD FOR PRODUCING A BIOLOGICALLY
; TITLE OF INVENTION: ACTIVE RECOMBINANT CYSTEINE-FREE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,231
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,709
; FILING DATE: 07-DEC-1994
; APPLICATION NUMBER: 07/838,857
; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: JP 024841
; FILING DATE: 19-FEB-1991
; APPLICATION NUMBER: JP 0271438
; FILING DATE: 18-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DAVID, RESNICK S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 41614-FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: 200291 STRE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-08-835-231-8

  Query Match             98.3%;  Score 91.4;  DB 2;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 3.6e-25;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 2
US-09-108-661-8
; Sequence 8, Application US/09108661
; Patent No. 6287806
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 6287806uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/108,661
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  528 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-09-108-661-8

  Query Match             98.3%;  Score 91.4;  DB 4;  Length 528;
  Best Local Similarity   98.9%;  Pred. No. 3.6e-25;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 3
US-08-784-582-57
; Sequence 57, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  955 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-784-582-57
```

```
  Query Match             98.3%;  Score 91.4;  DB 3;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 4.4e-25;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 4
US-08-784-582-60
; Sequence 60, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  955 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-784-582-60

  Query Match             98.3%;  Score 91.4;  DB 3;  Length 955;
  Best Local Similarity   98.9%;  Pred. No. 4.4e-25;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 5
US-08-784-582-72
; Sequence 72, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2356 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-784-582-72

  Query Match             98.3%;  Score 91.4;  DB 3;  Length 2356;
  Best Local Similarity   98.9%;  Pred. No. 6.2e-25;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1704 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 1763

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db    1764 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 1796
```

```
RESULT 6
US-08-835-231-7
; Sequence 7, Application US/08835231
; Patent No. 5861284
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 5861284uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/835,231
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  528 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-08-835-231-7

  Query Match             96.6%;  Score 89.8;  DB 2;  Length 528;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-24;
  Matches   91;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACGAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 7
```

```
US-09-108-661-7
; Sequence 7, Application US/09108661
; Patent No. 6287806
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 6287806uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/108,661
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  528 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Synthetic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic DNA
US-09-108-661-7

  Query Match             96.6%;  Score 89.8;  DB 4;  Length 528;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-24;
  Matches   91;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1 CATGCTGAAGGGACCTTTACGAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 8
US-08-784-582-55
```

```
; Sequence 55, Application US/08784582
; Patent No. 6110707
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Philippe A.
;    APPLICANT:  No. 6110707mington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Anice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Kruse, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,427
;      FILING DATE:  15-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/589,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  895 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-784-582-55

  Query Match             88.0%;  Score 81.8;  DB 3;  Length 895;
  Best Local Similarity   92.5%;  Pred. No. 1.7e-21;
  Matches   86;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     343 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 402

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     403 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 435


RESULT 9
US-08-835-231-17
; Sequence 17, Application US/08835231
; Patent No. 5861284
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 5861284uyuki
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/835,231
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  144 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  22...144
US-08-835-231-17

  Query Match             85.2%;  Score 79.2;  DB 2;  Length 144;
  Best Local Similarity   91.3%;  Pred. No. 8.3e-21;
  Matches   84;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| |||||| |||||||||| |||||||| |||||| |||||||||| ||||||||
Db      43 CATGCTGAAGGCACCTTTACCAGCGATGTAAGCTCTTATCTGGAAGGCCAGGCTGCCAAA 102

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGG 92
           |||||||||||||||||||||||||||||| ||
Db     103 GAATTCATTGCTTGGCTGGTGAAAGGCCGTGG 134


RESULT 10
US-09-108-661-17
; Sequence 17, Application US/09108661
; Patent No. 6287806
;  GENERAL INFORMATION:
;    APPLICANT:  NISHIMURA, Osamu
;    APPLICANT:  KURIYAMA, Masato
;    APPLICANT:  KOYAMA, No. 6287806uyuki
```

```
;    APPLICANT:  FUKUDA, Tsunehiko
;    TITLE OF INVENTION:  METHOD FOR PRODUCING A BIOLOGICALLY
;    TITLE OF INVENTION:  ACTIVE RECOMBINANT CYSTEINE-FREE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/108,661
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/350,709
;      FILING DATE:  07-DEC-1994
;      APPLICATION NUMBER:  07/838,857
;      FILING DATE:  18-FEB-1992
;      APPLICATION NUMBER:  JP 024841
;      FILING DATE:  19-FEB-1991
;      APPLICATION NUMBER:  JP 0271438
;      FILING DATE:  18-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  41614-FWC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  144 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  22...144
US-09-108-661-17

  Query Match             85.2%;  Score 79.2;  DB 4;  Length 144;
  Best Local Similarity   91.3%;  Pred. No. 8.3e-21;
  Matches   84;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
             |||| |||||| |||||||||| |||||||| |||||| |||||||||| ||||||||
Db      43 CATGCTGAAGGCACCTTTACCAGCGATGTAAGCTCTTATCTGGAAGGCCAGGCTGCCAAA 102

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGG 92
           |||||||||||||||||||||||||||||| ||
Db     103 GAATTCATTGCTTGGCTGGTGAAAGGCCGTGG 134


RESULT 11
US-08-829-876-22
; Sequence 22, Application US/08829876
; Patent No. 5962266
;  GENERAL INFORMATION:
;    APPLICANT:  White, Tyler R.
;    APPLICANT:  Damm, Deborah
;    APPLICANT:  Lesikar, David D.
;    APPLICANT:  McFadden, Kathleen
;    APPLICANT:  Garrick, Brett L.
;    TITLE OF INVENTION:  PROTEASE INHIBITOR PEPTIDES
;    NUMBER OF SEQUENCES:  228
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/829,876
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/436,555
;      FILING DATE:  08-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pelto, Don J.
;      REGISTRATION NUMBER:  33,754
;      REFERENCE/DOCKET NUMBER:  56324/106/SCNO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-829-876-22

  Query Match             53.5%;  Score 49.8;  DB 2;  Length 78;
  Best Local Similarity   82.6%;  Pred. No. 7.2e-10;
  Matches   57;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      24 TGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAA 83
           ||| ||   |||||| |||||||| |||||||| |||||||||| |||||| |||| ||
Db       1 TGACGTCTCTTCTTACTTGGAAGGTCAAGCTGCTAAGGAATTCATCGCTTGGTTGGTCAA 60

Qy      84 AGGCCGAGG 92
           |||  ||||
Db      61 AGGTAGAGG 69


RESULT 12
US-09-234-874A-22
; Sequence 22, Application US/09234874A
; Patent No. 6376648
;  GENERAL INFORMATION:
;       APPLICANT: White, Tyler R.
;                  Damm, Deborah
;                  Lesikar, David D.
;                  McFadden, Kathleen
;                  Garrick, Brett L.
;       TITLE OF INVENTION: PROTEASE INHIBITOR PEPTIDES
;       NUMBER OF SEQUENCES: 228
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Foley & Lardner
;            STREET: 3000 K Street, N.W., Suite 500
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: USA
;            ZIP: 20007-5109
```

```
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/234,874A
;             FILING DATE: 11-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/436,555
;             FILING DATE: 08-MAY-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Bent, Stephen
;             REGISTRATION NUMBER: 29,768
;             REFERENCE/DOCKET NUMBER: 056324/0106
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (202)672-5300
;             TELEFAX: (202)672-5399
;             TELEX: 904136
;    INFORMATION FOR SEQ ID NO: 22:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 78 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-234-874A-22

  Query Match            53.5%;  Score 49.8;  DB 4;  Length 78;
  Best Local Similarity  82.6%;  Pred. No. 7.2e-10;
  Matches   57;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      24 TGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAA 83
           ||| ||   |||||| |||||||| |||||||| |||||||||| |||||| |||| ||
Db       1 TGACGTCTCTTCTTACTTGGAAGGTCAAGCTGCTAAGGAATTCATCGCTTGGTTGGTCAA 60

Qy      84 AGGCCGAGG 92
           |||  ||||
Db      61 AGGTAGAGG 69


RESULT 13
US-08-811-028-43
; Sequence 43, Application US/08811028C
; Patent No. 5891671
; GENERAL INFORMATION:
;  APPLICANT: SUZUKI, Yuji
;  APPLICANT: MAGOTA, Koji
;  APPLICANT: MASUDA, Toyofumi
;  TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING
;  TITLE OF INVENTION: ENZYME
;  FILE REFERENCE: 001560-294
;  CURRENT APPLICATION NUMBER: US/08/811,028C
;  CURRENT FILING DATE: 1987-03-04
;  EARLIER APPLICATION NUMBER: JP 8-70906
;  EARLIER FILING DATE: 1996-03-04
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
;   OTHER INFORMATION: GLP-1
US-08-811-028-43

  Query Match            40.2%;  Score 37.4;  DB 2;  Length 57;
  Best Local Similarity  80.0%;  Pred. No. 2.9e-05;
  Matches   44;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       7 GAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGG 61
           ||||| |||||||||| ||||| || || ||| ||||||| || || || || |
Db       3 GAAGGTACCTTTACCAGCGATGTGAGCTCGTATCTGGAAGGTCAGGCGGCAAAAG 57


RESULT 14
US-08-811-028-44/c
; Sequence 44, Application US/08811028C
; Patent No. 5891671
; GENERAL INFORMATION:
;  APPLICANT: SUZUKI, Yuji
;  APPLICANT: MAGOTA, Koji
;  APPLICANT: MASUDA, Toyofumi
;  TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING
;  TITLE OF INVENTION: ENZYME
;  FILE REFERENCE: 001560-294
;  CURRENT APPLICATION NUMBER: US/08/811,028C
;  CURRENT FILING DATE: 1987-03-04
;  EARLIER APPLICATION NUMBER: JP 8-70906
;  EARLIER FILING DATE: 1996-03-04
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
;   OTHER INFORMATION: GLP-2
US-08-811-028-44

  Query Match            36.8%;  Score 34.2;  DB 2;  Length 48;
  Best Local Similarity  83.0%;  Pred. No. 0.00043;
  Matches   39;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGG 47
           ||||  ||||| |||||||||||| ||||| || || ||| |||||||
Db      48 CATGCGGAAGGTACCTTTACCAGCGATGTGAGCTCGTATCTGGAAGG 2


RESULT 15
US-08-811-028-46/c
; Sequence 46, Application US/08811028C
; Patent No. 5891671
; GENERAL INFORMATION:
;  APPLICANT: SUZUKI, Yuji
;  APPLICANT: MAGOTA, Koji
;  APPLICANT: MASUDA, Toyofumi
;  TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING
;  TITLE OF INVENTION: ENZYME
;  FILE REFERENCE: 001560-294
;  CURRENT APPLICATION NUMBER: US/08/811,028C
;  CURRENT FILING DATE: 1987-03-04
;  EARLIER APPLICATION NUMBER: JP 8-70906
;  EARLIER FILING DATE: 1996-03-04
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
;   OTHER INFORMATION: GLP-4
US-08-811-028-46

  Query Match            35.3%;  Score 32.8;  DB 2;  Length 53;
  Best Local Similarity  84.1%;  Pred. No. 0.0015;
  Matches   37;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      49 CAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAAGGCCGAGG 92
           || || || || |||||||| || ||||||||||||||||| ||
Db      53 CAGGCGGCAAAAGAATTCATCGCGTGGCTGGTGAAAGGCCGTGG 10
```

Search completed: February 14, 2003, 08:25:51
Job time : 49 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:57:09 ; Search time 1184 Seconds
(without alignments)
1272.112 Million cell updates/sec

Title: US-09-091-605-4
Perfect score: 93
Sequence: 1 CATGTTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	91.4	98.3	359	13	BI715164	BI715164 ic29g03.y
	2	91.4	98.3	382	13	BM313323	BM313323 ig82f07.y
	3	91.4	98.3	389	14	BQ632756	BQ632756 il27e09.y
	4	91.4	98.3	394	12	BG654963	BG654963 ib44a12.y
	5	91.4	98.3	419	14	BM836042	BM836042 K-EST0111
	6	91.4	98.3	427	13	BI466966	BI466966 ic17d08.y
	7	91.4	98.3	443	14	BQ271272	BQ271272 ik11g04.y
	8	91.4	98.3	451	13	BM503895	BM503895 ig97b05.y
c	9	91.4	98.3	451	14	BQ776591	BQ776591 il34g04.x
	10	91.4	98.3	458	12	BG656237	BG656237 ib38g04.y
	11	91.4	98.3	459	14	BQ286311	BQ286311 ik28e04.y
	12	91.4	98.3	461	13	BM312520	BM312520 ig75e12.y
	13	91.4	98.3	463	14	BQ271456	BQ271456 ik14b08.y
	14	91.4	98.3	466	14	BQ416911	BQ416911 ik39c06.y
	15	91.4	98.3	468	14	BQ271361	BQ271361 ik12h07.y
	16	91.4	98.3	471	12	BG655087	BG655087 ib45h03.y
	17	91.4	98.3	471	14	BQ269449	BQ269449 ik25g03.y
	18	91.4	98.3	472	14	BQ777016	BQ777016 il42c09.y
	19	91.4	98.3	472	14	BQ787174	BQ787174 il51c09.y
c	20	91.4	98.3	474	13	BI467273	BI467273 ic22c02.x
	21	91.4	98.3	474	14	BQ271282	BQ271282 ik11h07.y
	22	91.4	98.3	476	14	BQ777432	BQ777432 il48b03.y
	23	91.4	98.3	478	14	BQ272090	BQ272090 ik18b09.y
	24	91.4	98.3	479	13	BM312257	BM312257 ig41b09.y
	25	91.4	98.3	479	14	BQ271314	BQ271314 ik12c06.y
	26	91.4	98.3	481	13	BM313155	BM313155 ig80c02.y
	27	91.4	98.3	481	14	BQ776373	BQ776373 il31d12.y
	28	91.4	98.3	483	13	BI715416	BI715416 ic32f01.y
	29	91.4	98.3	483	14	BQ269264	BQ269264 ik23d04.y
	30	91.4	98.3	484	14	BQ271926	BQ271926 ik16b09.y
	31	91.4	98.3	484	14	BQ417049	BQ417049 ik41b03.y
	32	91.4	98.3	484	14	BQ776874	BQ776874 il34d03.y
	33	91.4	98.3	485	14	BQ631915	BQ631915 il24a04.y
	34	91.4	98.3	485	14	BQ777960	BQ777960 il39c04.y
	35	91.4	98.3	486	14	BQ778021	BQ778021 il40a04.y
	36	91.4	98.3	487	13	BM312448	BM312448 ig78f12.x
	37	91.4	98.3	487	14	BQ269028	BQ269028 ik20b03.y
	38	91.4	98.3	487	14	BQ269433	BQ269433 ik25e08.y
	39	91.4	98.3	490	13	BM509838	BM509838 ig94f02.y
	40	91.4	98.3	492	13	BM312759	BM312759 ig78g03.y
	41	91.4	98.3	494	13	BI791789	BI791789 ie04e11.y
	42	91.4	98.3	497	12	BG654541	BG654541 ib41h07.y
c	43	91.4	98.3	497	13	BI438781	BI438781 ic26f01.x
	44	91.4	98.3	497	13	BI792098	BI792098 ie06f11.y
	45	91.4	98.3	501	12	BE970164	BE970164 601680194

ALIGNMENTS

RESULT 1
BI715164
LOCUS BI715164 359 bp mRNA linear EST 19-SEP-2001
DEFINITION ic29g03.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BI715164
VERSION BI715164.1 GI:15690859
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Sceare,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 277.

FEATURES Location/Qualifiers
source 1. .359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT 103 a 74 c 92 g 90 t
ORIGIN

Query Match 98.3%; Score 91.4; DB 13; Length 359;
Best Local Similarity 98.9%; Pred. No. 1.6e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 185

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     186 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 218
```

RESULT 2
BM313323
LOCUS BM313323 382 bp mRNA linear EST 03-JAN-2002
DEFINITION ig82f07.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BM313323
VERSION BM313323.1 GI:18047668
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 382)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: ig82f07.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco.

FEATURES Location/Qualifiers
source 1. .382
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT 110 a 79 c 98 g 95 t
ORIGIN

Query Match 98.3%; Score 91.4; DB 13; Length 382;
Best Local Similarity 98.9%; Pred. No. 1.7e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 194

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     195 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 227
```

RESULT 3
BQ632756
LOCUS BQ632756 389 bp mRNA linear EST 02-JUL-2002
DEFINITION il27e09.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6031216 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BQ632756
VERSION BQ632756.1 GI:21684274
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 389)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: il27e09.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 361.

FEATURES Location/Qualifiers
source 1. .389
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:6031216"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

BASE COUNT 116 a 76 c 94 g 103 t
ORIGIN

Query Match 98.3%; Score 91.4; DB 14; Length 389;
Best Local Similarity 98.9%; Pred. No. 1.7e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 131

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     132 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 164
```

RESULT 4
BG654963
LOCUS BG654963 394 bp mRNA linear EST 05-JUL-2001
DEFINITION ib44a12.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BG654963
VERSION BG654963.1 GI:13792372
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 394)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 350.
FEATURES Location/Qualifiers
source 1. .394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

BASE COUNT 119 a 74 c 90 g 110 t 1 others
ORIGIN

Query Match 98.3%; Score 91.4; DB 12; Length 394;
Best Local Similarity 98.9%; Pred. No. 1.7e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 117

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     118 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 150
```

RESULT 5
BM836042
LOCUS BM836042 419 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0111434 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-18-G08 5', mRNA sequence.
ACCESSION BM836042
VERSION BM836042.1 GI:19192451
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 419)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 18 row: G column: 08
High quality sequence stop: 419.
FEATURES Location/Qualifiers
source 1. .419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S5SNU484s1-18-G08"
/clone_lib="S5SNU484s1"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F'"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tabacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method.

After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promotor as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original liberary and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformaion of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

BASE COUNT 122 a 82 c 94 g 121 t
ORIGIN

Query Match 98.3%; Score 91.4; DB 14; Length 419;
Best Local Similarity 98.9%; Pred. No. 1.7e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 89

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      90 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 122
```

RESULT 6
BI466966
LOCUS BI466966 427 bp mRNA linear EST 22-AUG-2001
DEFINITION ic17d08.y3 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BI466966
VERSION BI466966.1 GI:15279844
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 427)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: ic17d08.x3
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Trace considered overall poor quality
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

BASE COUNT 123 a 94 c 111 g 99 t
ORIGIN

Query Match 98.3%; Score 91.4; DB 13; Length 427;
Best Local Similarity 98.9%; Pred. No. 1.8e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     290 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 349

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     350 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 382
```

RESULT 7
BQ271272
LOCUS BQ271272 443 bp mRNA linear EST 07-MAY-2002
DEFINITION ik11g04.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5780910 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BQ271272
VERSION BQ271272.1 GI:20496338
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco.
FEATURES Location/Qualifiers
source 1. .443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE: 5780910"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University

School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."
BASE COUNT 132 a 99 c 110 g 101 t 1 others
ORIGIN

Query Match 98.3%; Score 91.4; DB 14; Length 443;
Best Local Similarity 98.9%; Pred. No. 1.8e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 335 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 394

Qy 61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
|||||||||||||||||||||||||||||||||
Db 395 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 427

RESULT 8
BM503895
LOCUS BM503895 451 bp mRNA linear EST 14-FEB-2002
DEFINITION ig97b05.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BM503895
VERSION BM503895.1 GI:18666121
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: ig97b05.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 440.
FEATURES Location/Qualifiers
source 1. .451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."
BASE COUNT 135 a 86 c 106 g 123 t 1 others
ORIGIN

Query Match 98.3%; Score 91.4; DB 13; Length 451;
Best Local Similarity 98.9%; Pred. No. 1.8e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 74 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 133

Qy 61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
|||||||||||||||||||||||||||||||||
Db 134 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 166

RESULT 9
BQ776591/c
LOCUS BQ776591 451 bp mRNA linear EST 26-JUL-2002
DEFINITION il34g04.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6032047 3' similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION BQ776591
VERSION BQ776591.1 GI:21985063
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco.
FEATURES Location/Qualifiers
source 1. .451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6032047"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."
BASE COUNT 132 a 96 c 82 g 141 t
ORIGIN

Query Match 98.3%; Score 91.4; DB 14; Length 451;
Best Local Similarity 98.9%; Pred. No. 1.8e-20;
Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    451 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 392

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    391 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 359


RESULT 10
BG656237
LOCUS       BG656237                458 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  ib38g04.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN
            P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION   BG656237
VERSION     BG656237.1  GI:13793646
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 452.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      134 a    106 c    111 g    107 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 12;  Length 458;
  Best Local Similarity   98.9%;  Pred. No. 1.8e-20;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    348 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 407

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    408 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 440


RESULT 11
BQ286311
LOCUS       BQ286311                459 bp    mRNA    linear   EST 14-MAY-2002
DEFINITION  ik28e04.y1 HR85 islet Homo sapiens cDNA clone IMAGE:5782351 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ286311
VERSION     BQ286311.1  GI:20655687
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 459)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 440.
FEATURES             Location/Qualifiers
     source          1. .459
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5782351"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      135 a    105 c    109 g    110 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 14;  Length 459;
  Best Local Similarity   98.9%;  Pred. No. 1.8e-20;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 425

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    426 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 458


RESULT 12
```

```
BM312520
LOCUS       BM312520                 461 bp    mRNA    linear   EST 03-JAN-2002
DEFINITION  ig75e12.y1 HR85 islet Homo sapiens cDNA 5' similar to SW:GLUC_HUMAN
            P01275 GLUCAGON PRECURSOR. [1] ;, mRNA sequence.
ACCESSION   BM312520
VERSION     BM312520.1  GI:18046865
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 401.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      128 a    102 c    118 g    113 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 13;  Length 461;
  Best Local Similarity   98.9%;  Pred. No. 1.8e-20;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 324

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     325 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 357


RESULT 13
BQ271456
LOCUS       BQ271456                 463 bp    mRNA    linear   EST 07-MAY-2002
DEFINITION  ik14b08.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5780703 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ271456
VERSION     BQ271456.1  GI:20496522
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 463)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ik14b08.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .463
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE: 5780703"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      135 a    107 c    115 g    106 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 14;  Length 463;
  Best Local Similarity   98.9%;  Pred. No. 1.8e-20;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 420

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     421 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 453


RESULT 14
BQ416911
LOCUS       BQ416911                 466 bp    mRNA    linear   EST 23-MAY-2002
DEFINITION  ik39c06.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5783410 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ416911
VERSION     BQ416911.1  GI:21122112
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 466)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ik39c06.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .466
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE: 5783410"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      139 a    108 c    113 g    106 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 14;  Length 466;
  Best Local Similarity   98.9%;  Pred. No. 1.8e-20;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 429

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    430 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 462


RESULT 15
BQ271361
LOCUS       BQ271361                468 bp    mRNA    linear   EST 07-MAY-2002
DEFINITION  ik12h07.y1 HR85 islet Homo sapiens cDNA clone IMAGE: 5780965 5'
            similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ;, mRNA
            sequence.
ACCESSION   BQ271361
VERSION     BQ271361.1  GI:20496427
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ik12h07.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE: 5780965"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      140 a    108 c    114 g    106 t
ORIGIN

  Query Match             98.3%;  Score 91.4;  DB 14;  Length 468;
  Best Local Similarity   98.9%;  Pred. No. 1.8e-20;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 429

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db    430 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 462


Search completed: February 14, 2003, 09:05:35
Job time : 1185 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 07:10:09 ; Search time 174.5 Seconds
(without alignments)
1200.205 Million cell updates/sec

Title: US-09-091-605-2
Perfect score: 93
Sequence: 1 CATGCTGAAGGGACCTTTAC.....GGCTGGTGAAAGGCCGAGGA 93

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	93	100.0	93	18	AAT77296	DNA encoding gluca
2	93	100.0	96	13	AAQ27605	DNA encoding gluca
3	93	100.0	955	18	AAT75672	Human preproglucag
4	93	100.0	955	18	AAT75673	Human mutated prep
5	93	100.0	955	21	AAC55763	Human preproglucag
6	93	100.0	955	21	AAC55765	Mutant human prepr
7	93	100.0	2356	21	AAC55775	Human growth hormo
8	93	100.0	3798	23	ABV25306	Human prostate exp
9	91.4	98.3	93	18	AAT77297	DNA encoding gluca
10	91.4	98.3	396	22	AAF30989	Prepro-somatostati
11	88	94.6	626	24	ABQ58859	Human colon cancer
12	83.4	89.7	895	18	AAT75669	Rat preproglucagon
13	83.4	89.7	895	21	AAC55762	Rat preproglucagon
14	83.4	89.7	1034	11	AAQ06255	Glucagon-like pept
15	83.4	89.7	1034	18	AAT73216	Rat prepro-glucago
16	83.4	89.7	1034	20	AAZ20678	Preproglucagon cod
17	81.8	88.0	387	21	AAA51462	PCpB-RR-GLIP (R26)
18	81.8	88.0	390	21	AAA51461	PCpB-RVR-GLIP (R26
19	81.8	88.0	390	21	AAA51464	PCpB-QAR-GLIP (R26
20	81.8	88.0	393	21	AAA51463	PCpB-EGGR-GLIP (R2
21	80.8	86.9	144	13	AAQ27607	GLP-I(7-37)-hbFGF
22	80.4	86.5	387	21	AAA51460	PCpB-AM-GLIP (R26)
23	74.8	80.4	95	22	AAF29136	Glucagon-like pept
24	72.2	77.6	1576	13	AAQ25339	Chicken glucagon c
25	68	73.1	627	20	AAX99247	Nucleotide sequenc
26	63.2	68.0	315	11	AAQ04767	Sequence encoding
27	61.6	66.2	395	21	AAA51465	PCpB-QAR-V8-GLIP f
28	61.6	66.2	398	21	AAA51466	PCpB-LVPR-V8-GLIP
29	58.6	63.0	110	16	AAQ91253	Glucagon like pept
30	58.6	63.0	112	16	AAQ91251	Glucagon like pept
31	58.6	63.0	384	16	AAQ91259	Glucagon like pept
32	58.2	62.6	516	17	AAT07347	pAK623 fragment.
33	57	61.3	279	17	AAT34870	Plasmid pBN4:GLP(7
34	55.4	59.6	174	17	AAT34869	Plasmid pBN4:GLP(7
35	54.6	58.7	1134	22	AAC86599	DNA encoding a Bac
36	54.6	58.7	1158	22	AAC86600	DNA encoding a Bac
37	49.8	53.5	78	21	AAZ57465	Kunitz protease in
38	49.8	53.5	78	22	AAS14247	Oligonucleotide 62
c 39	47.4	51.0	112	16	AAQ91254	Glucagon like pept
c 40	47	50.5	108	16	AAQ91252	Glucagon like pept
41	37.8	40.6	57	18	AAV02282	Oligonucleotide GL
42	37.8	40.6	65	24	ABN56946	Mouse spliced tran
c 43	35.8	38.5	48	18	AAV02283	Oligonucleotide GL
c 44	35	37.6	70	22	AAC86614	PCR primer used to
45	34	36.6	492	19	AAV33163	Heloderma suspectu

ALIGNMENTS

RESULT 1
AAT77296
ID AAT77296 standard; DNA; 93 BP.
XX
AC AAT77296;
XX
DT 14-JAN-1998 (first entry)
XX
DE DNA encoding glucagon-like peptide GLP-1(7-37) Ala8 analogue.
XX
KW Diabetes; non-insulin dependent diabetes mellitus; NIDDM;
KW insulin dependent diabetes mellitus; IDDM; gene therapy;
KW glucagon-like peptide; GLP; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT mat_peptide 1..93
FT /*tag= a
FT /product= Ala8_GLP-1(7-37)
XX
PN WO9729180-A1.
XX
PD 14-AUG-1997.
XX
PF 06-FEB-1997; 97WO-US01978.
XX
PR 23-FEB-1996; 96GB-0003847.
PR 06-FEB-1996; 96US-0012111.
XX
PA (ELIL) LILLY & CO ELI.
XX

```
PI   Borts TL,  Broderick CL,  Dimarchi RD,  Grinnell BW;
PI   Miller AR;
XX
DR   WPI; 1997-415336/38.
DR   P-PSDB; AAW24389.
XX
PT   Gene therapy of type I and type II diabetes - by in vivo expression
PT   of glucagon like peptide GLP-1(7-37) analogue
XX
PS   Claim 8; Page 24; 31pp; English.
XX
CC   This DNA sequence encodes an analogue of a glucagon-like peptide
CC   GLP-1(7-37). This peptide provides a means of delivering long term
CC   amounts of a GLP-1(7-37)-based protein, which is useful in treating type
CC   I and type II diabetes. A stable mammalian cell line, which  is
CC   immunologically isolated from the mammal's immune system, is transformed
CC   with a vector expressing a protein of the above sequence. This
CC   transformed cell line is then implanted into the individual needing
CC   treatment. Once implanted, the GLP-1(7-37) analogue, in conjunction with
CC   high serum glucose levels, causes pancreatic cells to produce insulin in
CC   non-insulin dependent diabetes mellitus (NIDDM) and delays gastric
CC   emptying in both NIDDM and insulin dependent diabetes mellitus (IDDM)
CC   patients. An expression vector coding for a protein of the above
CC   sequence can also be directly injected into the mammal, such that the
CC   vector is incorporated into a cell and secretes the protein. This method
CC   overcomes the problems of the short serum half life of GLP-1(7-37),
CC   allowing delivery of effective long term amounts for diabetes treatment.
XX
SQ   Sequence 93 BP; 24 A; 17 C; 28 G; 24 T; 0 other;

  Query Match             100.0%;  Score 93;  DB 18;  Length 93;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-23;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 2
AAQ27605
ID   AAQ27605 standard; DNA; 96 BP.
XX
AC   AAQ27605;
XX
DT   04-FEB-1993  (first entry)
XX
DE   DNA encoding glucagon-like peptide I (7-37).
XX
KW   Human parathyroid hormone production; osteoporosis; GLP-I (7-37);
KW   hypoparathyroidism; hypertension; insulinotropin; ss.
XX
OS   Synthetic.
XX
PN   EP499990-A.
XX
PD   26-AUG-1992.
XX
PF   15-FEB-1992;   92EP-0102543.
XX
PR   19-FEB-1991;   91JP-0024841.
PR   18-OCT-1991;   91JP-0271438.
PR   24-OCT-1991;   91JP-0277724.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Fukuda T,  Koyama N,  Kuriyama M,  Nishimura O;
XX
DR   WPI; 1992-286114/35.
XX
PT   Cysteine-free peptide prodn., e.g. human parathyroid hormone
PT   deriv. - by culturing transformant to produce a fusion protein
PT   comprising the cysteine-free peptide fused to a cysteine at its
PT   N-terminus where cleavage can occur
XX
PS   Disclosure; Page 5; 60pp; English.
XX
CC   The DNA codes for glucagon-like peptide I (7-37) [GLP-I (7-37)]
CC   (insulinotropin). PTH (1-34), it may be used in a method of culturing
CC   a transformant to produce a fusion protein comprising a cysteine-free
CC   peptide fused to a cysteine at its N-terminus where cleavage can occur.
CC   This method can be used to produce peptides which can be used as a
CC   pharmaceutical or industry in general. See also AAQ27603-Q27622.
XX
SQ   Sequence 96 BP; 24 A; 18 C; 29 G; 25 T; 0 other;

  Query Match             100.0%;  Score 93;  DB 13;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-23;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
          |||||||||||||||||||||||||||||||||
Db     61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 3
AAT75672
ID   AAT75672 standard; DNA; 955 BP.
XX
AC   AAT75672;
XX
DT   03-FEB-1998  (first entry)
XX
DE   Human preproglucagon cDNA.
XX
KW   Recombinant protein; expression; secretory cell line; human;
KW   glucagon; peptide hormone; amidation; insulinoma; RIN; rat;
KW   diabetes; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             27..569
FT                   /*tag=  a
XX
PN   WO9726321-A2.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997;   97WO-US00761.
XX
PR   15-OCT-1996;   96US-0028427.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (BETA-) BETAGENE INC.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Clark SA,  Halban PA,  Kruse F,  McGarry D,  Newgard CB;
PI   Normington KD, Quaade C,  Thigpen AE;
XX
DR   WPI; 1997-385326/35.
DR   P-PSDB; AAW22080.
XX
PT   Recombinant cell engineered to provide amylin to a mammal - useful
PT   to treat e.g. angiogenesis, anorexia, obesity, hypertension,
PT   osteoporosis etc.
XX
PS   Example 10; Page 283-284; 336pp; English.
```

```
XX
CC   This cDNA sequence includes a coding sequence for human
CC   preproglucagon (see AAW22080).  It was produced from pancreatic
CC   cDNA by PCR amplification (see AAT75670-71), and has been ligated
CC   into pNoTA/T7, generating pNoTAT7/h.glucagon.  A mutated human
CC   preproglucagon cDNA (see AAT75673) has also been produced.  Rat
CC   insulinoma RIN cell lines expressing preproglucagon demonstrated
CC   efficient amidation of a secreted, processed polypeptide.  The
CC   invention provides methods for production of heterologous
CC   polypeptides using recombinantly engineered cell lines.  Also
CC   described are methods of engineering cells for high level
CC   expression, methods of large-scale heterologous protein production,
CC   and methods for treatment of disease in vivo using viral delivery
CC   systems and recombinant cell lines.
XX
SQ   Sequence 955 BP; 301 A; 181 C; 203 G; 270 T; 0 other;

  Query Match             100.0%;  Score 93;  DB 18;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-23;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 4
AAT75673
ID   AAT75673 standard; DNA; 955 BP.
XX
AC   AAT75673;
XX
DT   03-FEB-1998  (first entry)
XX
DE   Human mutated preproglucagon cDNA.
XX
KW   Recombinant protein; expression; secretory cell line; human;
KW   glucagon; peptide hormone; amidation; insulinoma; RIN; rat;
KW   diabetes; gene therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             27..569
FT                   /*tag=  a
FT   mutation        181
FT                   /*tag=  b
FT                   /note= "C at position 181 creates an Ala-52 codon"
FT   mutation        204
FT                   /*tag=  b
FT                   /note= "T at position 204 creates an SpeI site"
XX
PN   WO9726321-A2.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997;   97WO-US00761.
XX
PR   15-OCT-1996;   96US-0028427.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (BETA-) BETAGENE INC.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Clark SA,  Halban PA,  Kruse F,  McGarry D,  Newgard CB;
PI   Normington KD, Quaade C,  Thigpen AE;
XX
DR   WPI; 1997-385326/35.
DR   P-PSDB; AAW22081.
XX
PT   Recombinant cell engineered to provide amylin to a mammal - useful
PT   to treat e.g. angiogenesis, anorexia, obesity, hypertension,
PT   osteoporosis etc.
XX
PS   Example 10; Page 285-286; 336pp; English.
XX
CC   This cDNA sequence includes a coding sequence for a mutated human
CC   preproglucagon (see AAW22081).  It was produced from plasmid
CC   pNoTAT7/h.glucagon (see AAT75672) using mutagenic primers.  2
CC   Mutations were created, one altering the native Arg-52 codon to an
CC   Ala codon, and one creating an SpeI site, with no change in the
CC   deduced amino acid sequence.  The mutated sequence has been used
CC   to construct plasmid pNiTAT7/mut.glucagon.  Rat insulinoma RIN cell
CC   lines expressing preproglucagon demonstrated efficient amidation of
CC   a secreted, processed polypeptide.  The invention provides methods
CC   for production of heterologous polypeptides using recombinantly
CC   engineered cell lines.  Also described are methods of engineering
CC   cells for high level expression, methods of large-scale
CC   heterologous protein production, and methods for treatment of
CC   disease in vivo using viral delivery systems and recombinant cell
CC   lines.
XX
SQ   Sequence 955 BP; 302 A; 180 C; 202 G; 271 T; 0 other;

  Query Match             100.0%;  Score 93;  DB 18;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-23;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 5
AAC55763
ID   AAC55763 standard; cDNA; 955 BP.
XX
AC   AAC55763;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human preproglucagon encoding cDNA.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (BETA-) BETAGENE INC.
XX
PI   Newgard CB,  Halban P,  Normington KD,  Thigpen AE,  Quaade C;
PI   Kruse F,  McGarry D,  Clark SA;
XX
DR   WPI; 2000-586352/55.
DR   P-PSDB; AAB26774.
XX
```

```
PT   Producing mammalian amylin, useful for regulating blood glucose and
PT   insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT   employing recombinantly engineered secretory cell lines -
XX
PS   Example 10; Column 175-178; 136pp; English.
XX
CC   This invention relates to a method for producing mammalian amylin. The
CC   method relies on the use of a recombinantly engineered secretory cell
CC   line. The method comprises:
CC   (a) providing a starting secretory cell that has a regulated secretory
CC   pathway;
CC   (b) introducing, into the starting secretory cell, an amylin-encoding
CC   gene operatively linked to a first promoter;
CC   (c) selecting a secretory cell of (b) that exhibits increased production
CC   of biologically active amylin as compared to the starting secretory
CC   cell; and (d) culturing the selected secretory cell.
CC   Amylin is an insulin modulator, and the method results in antidiabetic,
CC   hypotensive and osteopathic activity. The amylins produced are useful
CC   for regulating blood glucose levels, as well as in modulating the
CC   circulating levels of insulin in a mammal. The amylins produced maybe
CC   used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC   disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC   requiring amylin regulation. The invention includes cDNA and protein
CC   sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC   AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC   the invention for the production of mammalian amylin.
XX
SQ   Sequence 955 BP; 301 A; 181 C; 203 G; 270 T; 0 other;

  Query Match             100.0%;  Score 93;  DB 21;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-23;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 6
AAC55765
ID   AAC55765 standard; cDNA; 955 BP.
XX
AC   AAC55765;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Mutant human preproglucagon cDNA.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (BETA-) BETAGENE INC.
XX
PI   Newgard CB,  Halban P,  Normington KD,  Thigpen AE,  Quaade C;
PI   Kruse F,  McGarry D,  Clark SA;
XX
DR   WPI; 2000-586352/55.
DR   P-PSDB; AAB26775.
XX
PT   Producing mammalian amylin, useful for regulating blood glucose and
PT   insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT   employing recombinantly engineered secretory cell lines -
XX
PS   Example 10; Column 179-180; 136pp; English.
XX
CC   This invention relates to a method for producing mammalian amylin. The
CC   method relies on the use of a recombinantly engineered secretory cell
CC   line. The method comprises:
CC   (a) providing a starting secretory cell that has a regulated secretory
CC   pathway;
CC   (b) introducing, into the starting secretory cell, an amylin-encoding
CC   gene operatively linked to a first promoter;
CC   (c) selecting a secretory cell of (b) that exhibits increased production
CC   of biologically active amylin as compared to the starting secretory
CC   cell; and (d) culturing the selected secretory cell.
CC   Amylin is an insulin modulator, and the method results in antidiabetic,
CC   hypotensive and osteopathic activity. The amylins produced are useful
CC   for regulating blood glucose levels, as well as in modulating the
CC   circulating levels of insulin in a mammal. The amylins produced maybe
CC   used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC   disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC   requiring amylin regulation. The invention includes cDNA and protein
CC   sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC   AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC   the invention for the production of mammalian amylin.
XX
SQ   Sequence 955 BP; 302 A; 180 C; 202 G; 271 T; 0 other;

  Query Match             100.0%;  Score 93;  DB 21;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-23;
  Matches   93;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 377

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     378 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 410


RESULT 7
AAC55775
ID   AAC55775 standard; DNA; 2356 BP.
XX
AC   AAC55775;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Human growth hormone and mutated proglucagon fusion DNA sequence.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Homo sapiens.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (BETA-) BETAGENE INC.
XX
```

PI Newgard CB, Halban P, Normington KD, Thigpen AE, Quaade C;
PI Kruse F, McGarry D, Clark SA;
XX
DR WPI; 2000-586352/55.
DR P-PSDB; AAB26777.
XX
PT Producing mammalian amylin, useful for regulating blood glucose and
PT insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT employing recombinantly engineered secretory cell lines -
XX
PS Example 12; Column 187-190; 136pp; English.
XX
CC This invention relates to a method for producing mammalian amylin. The
CC method relies on the use of a recombinantly engineered secretory cell
CC line. The method comprises:
CC (a) providing a starting secretory cell that has a regulated secretory
CC pathway;
CC (b) introducing, into the starting secretory cell, an amylin-encoding
CC gene operatively linked to a first promoter;
CC (c) selecting a secretory cell of (b) that exhibits increased production
CC of biologically active amylin as compared to the starting secretory
CC cell; and (d) culturing the selected secretory cell.
CC Amylin is an insulin modulator, and the method results in antidiabetic,
CC hypotensive and osteopathic activity. The amylins produced are useful
CC for regulating blood glucose levels, as well as in modulating the
CC circulating levels of insulin in a mammal. The amylins produced maybe
CC used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC requiring amylin regulation. The invention includes cDNA and protein
CC sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC the invention for the production of mammalian amylin.
XX
SQ Sequence 2356 BP; 614 A; 600 C; 581 G; 561 T; 0 other;

Query Match 100.0%; Score 93; DB 21; Length 2356;
Best Local Similarity 100.0%; Pred. No. 4.1e-23;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1704 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 1763

Qy 61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
|||||||||||||||||||||||||||||||||
Db 1764 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 1796

RESULT 8
ABV25306
ID ABV25306 standard; cDNA; 3798 BP.
XX
AC ABV25306;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 25297.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 4978-4979; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 3798 BP; 1166 A; 563 C; 616 G; 1364 T; 89 other;

Query Match 100.0%; Score 93; DB 23; Length 3798;
Best Local Similarity 100.0%; Pred. No. 4.8e-23;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3411 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 3470

Qy 61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
|||||||||||||||||||||||||||||||||
Db 3471 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 3503

RESULT 9
AAT77297
ID AAT77297 standard; DNA; 93 BP.
XX
AC AAT77297;
XX
DT 14-JAN-1998 (first entry)
XX
DE DNA encoding glucagon-like peptide GLP-1(7-37) Val8 analogue.
XX
KW Diabetes; non-insulin dependent diabetes mellitus; NIDDM;
KW insulin dependent diabetes mellitus; IDDM; gene therapy;
KW glucagon-like peptide; GLP; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT mat_peptide 1..93
FT /*tag= a
FT /product= Val8_GLP-1(7-37)
XX
PN WO9729180-A1.
XX
PD 14-AUG-1997.
XX
PF 06-FEB-1997; 97WO-US01978.
XX

```
PR   23-FEB-1996;   96GB-0003847.
PR   06-FEB-1996;   96US-0012111.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Borts TL,  Broderick CL,  Dimarchi RD,  Grinnell BW;
PI   Miller AR;
XX
DR   WPI; 1997-415336/38.
DR   P-PSDB; AAW24390.
XX
PT   Gene therapy of type I and type II diabetes - by in vivo expression
PT   of glucagon like peptide GLP-1(7-37) analogue
XX
PS   Claim 8; Page 24; 31pp; English.
XX
CC   This DNA sequence encodes an analogue of a glucagon-like peptide
CC   GLP-1(7-37). This peptide provides a means of delivering long term
CC   amounts of a GLP-1(7-37)-based protein, which is useful in treating type
CC   I and type II diabetes. A stable mammalian cell line, which  is
CC   immunologically isolated from the mammal's immune system, is transformed
CC   with a vector expressing a protein of the above sequence. This
CC   transformed cell line is then implanted into the individual needing
CC   treatment. Once implanted, the GLP-1(7-37) analogue, in conjunction with
CC   high serum glucose levels, causes pancreatic cells to produce insulin in
CC   non-insulin dependent diabetes mellitus (NIDDM) and delays gastric
CC   emptying in both NIDDM and insulin dependent diabetes mellitus (IDDM)
CC   patients. An expression vector coding for a protein of the above
CC   sequence can also be directly injected into the mammal, such that the
CC   vector is incorporated into a cell and secretes the protein. This method
CC   overcomes the problems of the short serum half life of GLP-1(7-37),
CC   allowing delivery of effective long term amounts for diabetes treatment.
XX
SQ   Sequence 93 BP; 24 A; 16 C; 28 G; 25 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 18;  Length 93;
  Best Local Similarity   98.9%;  Pred. No. 6e-23;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATGTTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93


RESULT 10
AAF30989
ID   AAF30989 standard; DNA; 396 BP.
XX
AC   AAF30989;
XX
DT   23-JUL-2001  (first entry)
XX
DE   Prepro-somatostatin/GLP-1 coding sequence of pXIT-39.
XX
KW   Somatostatin; glucagon-like peptide 1; GLP-1; antidiabetic;
KW   drug delivery; diabetes; gene therapy; pXIT-38; ss.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..390
FT                   /*tag= a
FT                   /product= "prepro-somatostatin/cleavage site/GLP-1
FT                            fusion"
XX
PN   WO200136643-A1.
XX
PD   25-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US31634.
XX
PR   19-NOV-1999;   99US-0166508.
XX
PA   (TRAN-) TRANSKARYOTIC THERAPIES INC.
XX
PI   Treco DA,  Concino MF,  Duguay SJ;
XX
DR   WPI; 2001-355636/37.
XX
PT   New nucleic acid constructs useful for transforming cells useful as a
PT   drug delivery vehicle -
XX
PS   Example 1; Fig 3; 89pp; English.
XX
CC   The present sequence is that of a DNA sequence encoding a
CC   prepro-somatostatin-glucagon-like peptide 1 (GLP-1) fusion protein.
CC   The coding region is flanked by a 5' BamHI site and a 3' XhoI site.
CC   A multibasic cleavage site separates the prepro-somatostatin and
CC   GLP-1 moieties of the fusion protein.  The DNA sequence in plasmid
CC   pXIT-39 was designed for the stable, constitutive expression of
CC   GLP-1 (7-37) in human cells.  This is an example of nucleic acid
CC   constructs of the invention designed for the expression of small
CC   peptides.  The small peptides are especially therapeutic peptides
CC   such as small hormones and antidiabetic peptides such as GLP-1,
CC   exedin-4 and gastric inhibitory polypeptide.  Claimed methods
CC   of treating a subject having diabetes involve administering the
CC   nucleic acid construct or a cell capable of expressing the small
CC   peptide.  Transfected primary or secondary cells or cell strains
CC   have wide applicability as vehicles or delivery systems for
CC   therapeutic proteins.  By controlling the number of cells introduced
CC   into an individual, one can control the amount of the protein
CC   delivered to the body.  In addition, in some cases, it is possible
CC   to remove the transfected cells if there is no longer a need for
CC   the product.  Human fibroblasts transfected with pXIT-39 secreted
CC   human GLP-1(7-37) and GLP-1(3-36) (see AAB82335-36).
XX
SQ   Sequence 396 BP; 80 A; 112 C; 121 G; 83 T; 0 other;

  Query Match             98.3%;  Score 91.4;  DB 22;  Length 396;
  Best Local Similarity   98.9%;  Pred. No. 9.2e-23;
  Matches   92;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     286 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 345

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||| ||
Db     346 GAATTCATTGCTTGGCTGGTGAAAGGCCGACGA 378


RESULT 11
ABQ58859
ID   ABQ58859 standard; cDNA; 626 BP.
XX
AC   ABQ58859;
XX
DT   02-AUG-2002  (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:2554.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200229086-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30732.
```

```
XX
PR   02-OCT-2000; 2000US-237271P.
XX
PA   (FARB ) BAYER CORP.
XX
PI   Burgess C,  Astle JH,  Carroll E,  Catino TJ,  Dwivedi P,  Molino GA;
PI   Thiaglingam A,  Lewis ME;
XX
DR   WPI; 2002-426115/45.
XX
PT   New isolated nucleic acid that is differentially expressed in cancer
PT   tissues useful for determining the presence of colon cancer in a cell
PT   or tissue type, and in antisense therapy  -
XX
PS   Claim 1; Fig 1; 796pp; English.
XX
CC   ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC   expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC   encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC   used in antisense therapy. An antibody immunoreactive with a polypeptide
CC   encoded by (I) is useful for detecting cancer in a patient sample, and
CC   for detecting the presence or absence of a polynucleotide encoded by a
CC   nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC   from (I) can be used for determining the presence of a nucleic acid which
CC   hybridises to (I), and for determining the phenotype of cells in a sample
CC   of cells from a patient. (I) is useful for determining the presence of
CC   colon cancer in a cell or tissue type, for determining the presence or
CC   state of other type of cancer, in antisense therapy, to generate
CC   macroarrays on a solid surface, to identify a chromosome on which the
CC   corresponding gene resides, and in tissue profiling, forensics, genetic
CC   analysis, mapping and diagnostic applications. (I) can be used to raise
CC   antibodies, and to screen for peptide analogues and antagonists.
XX
SQ   Sequence 626 BP; 170 A; 134 C; 143 G; 159 T; 20 other;

  Query Match             94.6%;  Score 88;  DB 24;  Length 626;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-21;
  Matches   88;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 312

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCC 88
           ||||||||||||||||||||||||||||
Db     313 GAATTCATTGCTTGGCTGGTGAAAGGCC 340


RESULT 12
AAT75669
ID   AAT75669 standard; DNA; 895 BP.
XX
AC   AAT75669;
XX
DT   03-FEB-1998  (first entry)
XX
DE   Rat preproglucagon cDNA.
XX
KW   Recombinant protein; expression; secretory cell line; rat;
KW   glucagon; peptide hormone; amidation; insulinoma; RIN; diabetes;
KW   gene therapy; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             52..594
FT                   /*tag=  a
XX
PN   WO9726321-A2.
XX
PD   24-JUL-1997.
XX
PF   17-JAN-1997;   97WO-US00761.
XX
PR   15-OCT-1996;   96US-0028427.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (BETA-) BETAGENE INC.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Clark SA,  Halban PA,  Kruse F,  McGarry D,  Newgard CB;
PI   Normington KD, Quaade C,  Thigpen AE;
XX
DR   WPI; 1997-385326/35.
DR   P-PSDB; AAW22079.
XX
PT   Recombinant cell engineered to provide amylin to a mammal - useful
PT   to treat e.g. angiogenesis, anorexia, obesity, hypertension,
PT   osteoporosis etc.
XX
PS   Example 10; Page 281-282; 336pp; English.
XX
CC   This cDNA sequence includes a coding sequence for rat
CC   preproglucagon (see AAW22079).  It was produced from pancreatic
CC   cDNA by PCR amplification (see AAT75667-68), and has been ligated
CC   into pNoTA/T7, generating pNoTAT7/r.Glucagon.  Rat insulinoma RIN
CC   cell lines expressing preproglucagon demonstrated efficient
CC   amidation of a secreted, processed polypeptide.  The invention
CC   provides methods for production of heterologous polypeptides using
CC   recombinantly engineered cell lines.  Also described are methods of
CC   engineering cells for high level expression, methods of large-scale
CC   heterologous protein production, and methods for treatment of
CC   disease in vivo using viral delivery systems and recombinant cell
CC   lines.
XX
SQ   Sequence 895 BP; 266 A; 211 C; 200 G; 218 T; 0 other;

  Query Match             89.7%;  Score 83.4;  DB 18;  Length 895;
  Best Local Similarity   93.5%;  Pred. No. 8.2e-20;
  Matches   87;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     343 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 402

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     403 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 435


RESULT 13
AAC55762
ID   AAC55762 standard; cDNA; 895 BP.
XX
AC   AAC55762;
XX
DT   17-JAN-2001  (first entry)
XX
DE   Rat preproglucagon encoding cDNA.
XX
KW   Amylin; production; secretory cell; blood glucose level regulation;
KW   diabetes mellitus; hypoglycaemia; osteoporosis; Paget's disease;
KW   hypercalcaemia; obesity; hypertension; ss.
XX
OS   Rattus sp.
XX
PN   US6110707-A.
XX
PD   29-AUG-2000.
XX
PF   17-JAN-1997;   97US-0784582.
XX
PR   11-OCT-1996;   96US-0028279.
PR   19-JAN-1996;   96US-0589028.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
```

```
PA   (BETA-) BETAGENE INC.
XX
PI   Newgard CB,  Halban P,  Normington KD,  Thigpen AE,  Quaade C;
PI   Kruse F,  McGarry D,  Clark SA;
XX
DR   WPI; 2000-586352/55.
DR   P-PSDB; AAB26773.
XX
PT   Producing mammalian amylin, useful for regulating blood glucose and
PT   insulin levels, e.g. for treating diabetes mellitus or hypoglycemia, by
PT   employing recombinantly engineered secretory cell lines -
XX
PS   Example 10; Column 173-174; 136pp; English.
XX
CC   This invention relates to a method for producing mammalian amylin. The
CC   method relies on the use of a recombinantly engineered secretory cell
CC   line. The method comprises:
CC   (a) providing a starting secretory cell that has a regulated secretory
CC   pathway;
CC   (b) introducing, into the starting secretory cell, an amylin-encoding
CC   gene operatively linked to a first promoter;
CC   (c) selecting a secretory cell of (b) that exhibits increased production
CC   of biologically active amylin as compared to the starting secretory
CC   cell; and (d) culturing the selected secretory cell.
CC   Amylin is an insulin modulator, and the method results in antidiabetic,
CC   hypotensive and osteopathic activity. The amylins produced are useful
CC   for regulating blood glucose levels, as well as in modulating the
CC   circulating levels of insulin in a mammal. The amylins produced maybe
CC   used in treating diabetes mellitus, hypoglycaemia, osteoporosis, Paget's
CC   disease, hypercalcaemia, obesity, hypertension, or any other disorder
CC   requiring amylin regulation. The invention includes cDNA and protein
CC   sequences (AAC55760 and AAB26771) representing human amylin. Sequences
CC   AAC55716-C55681 and AAB26765-B26777 are used in examples of the method of
CC   the invention for the production of mammalian amylin.
XX
SQ   Sequence 895 BP; 266 A; 212 C; 199 G; 218 T; 0 other;

  Query Match             89.7%;  Score 83.4;  DB 21;  Length 895;
  Best Local Similarity   93.5%;  Pred. No. 8.2e-20;
  Matches   87;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     343 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 402

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     403 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 435


RESULT 14
AAQ06255
ID   AAQ06255 standard; DNA; 1034 BP.
XX
AC   AAQ06255;
XX
DT   29-JAN-1991  (first entry)
XX
DE   Glucagon-like peptide, GLP-1 (7-37).
XX
KW   Insulin; diabetes mellitus; insulinotropic: pancreatic beta cells.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             61..603
FT                   /*tag=  a
FT   sig_peptide     61..120
FT                   /*tag=  b
XX
PN   WO9011296-A.
XX
PD   04-OCT-1990.
XX
PF   20-MAR-1989;   89WO-US01121.
XX
PR   20-MAR-1989;   89WO-US01121.
XX
PA   (GEHO-) GEN HOSPITAL CORP.
XX
PI   Habener JF;
XX
DR   WPI; 1990-320226/42.
XX
PT   New glucagon-like peptide (GLP-1) - having insulin
PT   formation-stimulating activity and useful in treating diabetes
PT   mellitus.
XX
PS   Claim 6; Page 39; 52pp; English.
XX
CC   The peptide has insulinotropic activity specifically for pancreatic
CC   beta cells. The peptide is derived from glucagon which, after
CC   synthesis is cleaved into three peptides: glucgon, glucagon-like
CC   peptide 1 (GLP-1) and GLP-2.  GLP-1 has 37 AAs in its unprocessed
CC   form and is unable to mediate the induction of insulin biosynthesis.
CC   It is, however, naturally converted to a 31 AA-long peptide having
CC   AAs 7-37 of GLP-1.  Preferred derivs. have an H2 gp at the
CC   N-terminal and an OH, OM, or NR'R'' gp at the C-terminal where M= a
CC   cation or lower alkyl gp., and R' and R''= H or a lower alkyl gp.
CC   Prepns. contg. the peptide or derivs. are useful in the study of
CC   the pathogenesis of maturity onset of diabetes mellitus and also in
CC   therapy.
CC   See also AAR07398.
XX
SQ   Sequence 1034 BP; 315 A; 238 C; 226 G; 255 T; 0 other;

  Query Match             89.7%;  Score 83.4;  DB 11;  Length 1034;
  Best Local Similarity   93.5%;  Pred. No. 8.6e-20;
  Matches   87;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     352 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 411

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     412 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 444


RESULT 15
AAT73216
ID   AAT73216 standard; DNA; 1034 BP.
XX
AC   AAT73216;
XX
DT   01-OCT-1997  (first entry)
XX
DE   Rat prepro-glucagon DNA.
XX
KW   Glucagon-like peptide-1(7-36); GLP-1; insulin secretagogue;
KW   insulinotropic hormone; type II diabetes mellitus; therapy; ss.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             61..603
FT                   /*tag=  a
FT   sig_peptide     61..120
FT                   /*tag=  b
FT   mat_peptide     121..600
FT                   /*tag=  c
XX
PN   US5614492-A.
XX
PD   25-MAR-1997.
XX
```

PF 05-MAY-1986; 86US-0859928.
XX
PR 05-SEP-1991; 91US-0756215.
PR 05-MAY-1986; 86US-0859928.
PR 26-JAN-1988; 88US-0148517.
PR 01-JUN-1990; 90US-0532111.
PR 23-NOV-1993; 93US-0156800.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Habener JF;
XX
DR WPI; 1997-201513/18.
DR P-PSDB; AAW16384.
XX
PT Glucagon-like peptide-1 fragment comprising amino acids 7-36 -
PT useful for enhancing insulin production in pancreatic islet cells,
PT especially for treating type II diabetes mellitus
XX
PS Disclosure; Fig 1-1A; 37pp; English.
XX
CC A DNA molecule (AAT73216) codes for rat preproglucagon (AAW16384), a
CC polypeptide that is processed by proteolytic cleavage to glucagon,
CC glucagon-like peptide-1 (GLP-1) and GLP-2. A peptide fragment
CC of GLP-1, GLP-1(7-36) (see also AAW16363), has been shown to possess
CC hormonal activity and can be used as an insulin secretagogue and
CC for the treatment of type II diabetes mellitus.
XX
SQ Sequence 1034 BP; 316 A; 237 C; 226 G; 255 T; 0 other;

Query Match 89.7%; Score 83.4; DB 18; Length 1034;
Best Local Similarity 93.5%; Pred. No. 8.6e-20;
Matches 87; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy       1 CATGCTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAG 60
           |||||||||||||||||||||||||||||| |||||||| ||||| ||||| || || |||
Db     352 CATGCTGAAGGGACCTTTACCAGTGATGTGAGTTCTTACTTGGAGGGCCAGGCAGCAAAG 411

Qy      61 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 93
           |||||||||||||||||||||||||||||||||
Db     412 GAATTCATTGCTTGGCTGGTGAAAGGCCGAGGA 444
```

Search completed: February 14, 2003, 08:03:03
Job time : 176.5 secs